

171/487

GI-19745307 264 --- **TVV** **HEKPIIYFKLYRQLPGEKEVAVDDAELN** **QINS**
 ORF84 **WO 2006/078318** **TVV** **HEKPIIYFKLYRQLPGEKEVAVDDAELN** **QINSEGO** PCT/US2005/027239
 GI-28810263 264 **TVV** **HEKPIIYFKLYRQLPGEKEVAVDDAELN** **QINSEGO**
 GI-21909640 150 **TVV** **HEKPIIYFKLYRQLPGEKEVAVDDAELN** **QINSEGO**
 GI-19224141 601 **PHSS** **VRVEANKEVTI** **LNHN** **ETLTF** **SGKNIWEN** **RPDOR** **PAKIQV** **OLL** **LNQ** **GKMPNQIQEV**

GI-19745307 305 --- **Q** **EISVTW** **TNQLVT** **DERGM** **AYIYSV** **KEVDK** **NGELLEPKD** **YIKRED**
 ORF84 305 --- **Q** **EISVTW** **TNQLVT** **DERGM** **AYIYSV** **KEVDK** **NGELLEPKD** **YIKRED**
 GI-28810263 305 --- **Q** **EISVTW** **TNQLVT** **DERGM** **AYIYSV** **KEVDK** **NGELLEPKD** **YIKRED**
 GI-21909640 191 --- **Q** **EISVTW** **TNQLVT** **DERGM** **AYIYSV** **KEVDK** **NGELLEPKD** **YIKRED**
 GI-19224141 661 **TKD** **NS** **SYH** **ERD** **LPKY** **DA** **NOEY** **KYSV** **EEV** **VPDGYKVS** **YLGND** **IFNT** **RETE** **FV** **EQNNF**

GI-19745307 349 **GLTV** **TNTYVK** **PTSG** **HYDIEV** **TFTG** **NGHIDIT** **EDTTP** **DIVS** **GENQMK**
 ORF84 349 **GLTV** **TNTYVK** **PTSG** **HYDIEV** **TFTG** **NGHIDIT** **EDTTP** **DIVS** **GENQMK**
 GI-28810263 349 **GLTV** **TNTYVK** **PTSG** **HYDIEV** **TFTG** **NGHIDIT** **EDTTP** **DIVS** **GENQMK**
 GI-21909640 235 **GLTV** **TNTYVK** **PTSG** **HYDIEV** **TFTG** **NGHIDIT** **EDTTP** **DIVS** **GENQMK**
 GI-19224141 721 **N** **EFG** **NAE** **LG** **QSG** **SKI** **I** **DE** **TLT** **S** **FKG** **KI** **W** **K** **N** **D** **T** **A** **B** **R** **P** **Q** **A** **I** **Q** **V** **L** **Y** **A** **D** **G** **V** **E** **G** **O** **T** **K**

GI-19745307 394 **Q** **IEG** **EDS** **KPIDEV** **T** **ENNLIE** **F** **G** **K** **N** **T** **M** **P** **G** **E** **E**
 ORF84 394 **Q** **IEG** **EDS** **KPIDEV** **T** **ENNLIE** **F** **G** **K** **N** **T** **M** **P** **G** **E** **E**
 GI-28810263 394 **Q** **IEG** **EDS** **KPIDEV** **T** **ENNLIE** **F** **G** **K** **N** **T** **M** **P** **G** **E** **E**
 GI-21909640 280 **Q** **IEG** **EDS** **KPIDEV** **T** **ENNLIE** **F** **G** **K** **N** **T** **M** **P** **G** **E** **E**
 GI-19224141 781 **R** **I** **S** **G** **S** **G** **N** **E** **W** **S** **F** **E** **F** **K** **N** **L** **K** **K** **Y** **N** **G** **T** **G** **N** **D** **I** **I** **Y** **S** **V** **K** **E** **V** **T** **P** **T** **G** **D** **V** **T** **Y** **S** **A** **N** **D** **I** **I** **N** **T** **K** **E** **V** **T** **O** **A** **G**

GI-19745307 424 --- **D** **G** **T** **N** **S** **N** **K** **Y** **E** **E** **V** **E** **D** **S** **R** **P** **V** **D** **T** **L** **S** **G** **L** **S** **S** **E** **Q** **G** **S** **G** **D** **M** **T** **I** **E** **E** **D** **S** **A** **T** **H** **I** **K** **F** **S** **K** **R** **D**
 ORF84 424 --- **D** **G** **T** **N** **S** **N** **K** **Y** **E** **E** **V** **E** **D** **S** **R** **P** **V** **D** **T** **L** **S** **G** **L** **S** **S** **E** **Q** **G** **S** **G** **D** **M** **T** **I** **E** **E** **D** **S** **A** **T** **H** **I** **K** **F** **S** **K** **R** **D**
 GI-28810263 424 --- **D** **G** **T** **N** **S** **N** **K** **Y** **E** **E** **V** **E** **D** **S** **R** **P** **V** **D** **T** **L** **S** **G** **L** **S** **S** **E** **Q** **G** **S** **G** **D** **M** **T** **I** **E** **E** **D** **S** **A** **T** **H** **I** **K** **F** **S** **K** **R** **D**
 GI-21909640 310 --- **D** **G** **T** **N** **S** **N** **K** **Y** **E** **E** **V** **E** **D** **S** **R** **P** **V** **D** **T** **L** **S** **G** **L** **S** **S** **E** **Q** **G** **S** **G** **D** **M** **T** **I** **E** **E** **D** **S** **A** **T** **H** **I** **K** **F** **S** **K** **R** **D**
 GI-19224141 841 **PKLEI** **ET** **LP** **L** **E** **S** **G** **A** **S** **G** **G** **T** **T** **V** **E** **D** **S** **R** **P** **V** **D** **T** **L** **S** **G** **L** **S** **S** **E** **Q** **G** **S** **G** **D** **M** **T** **I** **E** **E** **D** **S** **A** **T** **H** **I** **K** **F** **S** **K** **R** **D**

GI-19745307 473 **IDG** **K** **E** **L** **A** **G** **A** **T** **M** **E** **L** **R** **D** **S** **S** **G** **N** **T** **I** **S** **T** **W** **I** **S** **D** **G** **Q** **V** **K** **D** **F** **Y** **L** **M** **P** **G** **K** **Y** **T** **F** **V** **E** **T** **A** **A** **P** **D** **G** **Y** **E** **I** **A** **T** **A** **I** **T** **F** **T**
 ORF84 473 **IDG** **K** **E** **L** **A** **G** **A** **T** **M** **E** **L** **R** **D** **S** **S** **G** **N** **T** **I** **S** **T** **W** **I** **S** **D** **G** **Q** **V** **K** **D** **F** **Y** **L** **M** **P** **G** **K** **Y** **T** **F** **V** **E** **T** **A** **A** **P** **D** **G** **Y** **E** **I** **A** **T** **A** **I** **T** **F** **T**
 GI-28810263 473 **IDG** **K** **E** **L** **A** **G** **A** **T** **M** **E** **L** **R** **D** **S** **S** **G** **N** **T** **I** **S** **T** **W** **I** **S** **D** **G** **Q** **V** **K** **D** **F** **Y** **L** **M** **P** **G** **K** **Y** **T** **F** **V** **E** **T** **A** **A** **P** **D** **G** **Y** **E** **I** **A** **T** **A** **I** **T** **F** **T**
 GI-21909640 359 **IDG** **K** **E** **L** **A** **G** **A** **T** **M** **E** **L** **R** **D** **S** **S** **G** **N** **T** **I** **S** **T** **W** **I** **S** **D** **G** **Q** **V** **K** **D** **F** **Y** **L** **M** **P** **G** **K** **Y** **T** **F** **V** **E** **T** **A** **A** **P** **D** **G** **Y** **E** **I** **A** **T** **A** **I** **T** **F** **T**
 GI-19224141 901 **IDG** **K** **E** **L** **A** **G** **A** **T** **M** **E** **L** **R** **D** **S** **S** **G** **N** **T** **I** **S** **T** **W** **I** **S** **D** **G** **Q** **V** **K** **D** **F** **Y** **L** **M** **P** **G** **K** **Y** **T** **F** **V** **E** **T** **A** **A** **P** **D** **G** **Y** **E** **I** **A** **T** **A** **I** **T** **F** **T**

GI-19745307 533 **V** **N** **E** **Q** **Q** **V** **T** **V** **N** **G** **K** **A** **T** **G** **D** **A** **H** **I** **V** **M** **D** **A** **Y** **K** **P** **T** **K** **S** **G** **Q** **V** **I** **D** **I** **E** **E** **K** **L** **F** **D** **E** **Q** **G** **H** **S** **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S**
 ORF84 533 **V** **N** **E** **Q** **Q** **V** **T** **V** **N** **G** **K** **A** **T** **G** **D** **A** **H** **I** **V** **M** **D** **A** **Y** **K** **P** **T** **K** **S** **G** **Q** **V** **I** **D** **I** **E** **E** **K** **L** **F** **D** **E** **Q** **G** **H** **S** **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S**
 GI-28810263 533 **V** **N** **E** **Q** **Q** **V** **T** **V** **N** **G** **K** **A** **T** **G** **D** **A** **H** **I** **V** **M** **D** **A** **Y** **K** **P** **T** **K** **S** **G** **Q** **V** **I** **D** **I** **E** **E** **K** **L** **F** **D** **E** **Q** **G** **H** **S** **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S**
 GI-21909640 419 **V** **N** **E** **Q** **Q** **V** **T** **V** **N** **G** **K** **A** **T** **G** **D** **A** **H** **I** **V** **M** **D** **A** **Y** **K** **P** **T** **K** **S** **G** **Q** **V** **I** **D** **I** **E** **E** **K** **L** **F** **D** **E** **Q** **G** **H** **S** **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S**
 GI-19224141 961 **V** **N** **E** **Q** **Q** **V** **T** **V** **N** **G** **K** **A** **T** **G** **D** **A** **H** **I** **V** **M** **D** **A** **Y** **K** **P** **T** **K** **S** **G** **Q** **V** **I** **D** **I** **E** **E** **K** **L** **F** **D** **E** **Q** **G** **H** **S** **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S**

GI-19745307 593 **SD** **V** **I** **I** **G** **Q** **G** **SD** **V** **I** **I** **G** **Q** **G** **E** **V** **D** **T** **T** **E** **D** **T** **Q** **S** **G** **M** **T** **G** **H** **S**
 ORF84 593 **SD** **V** **I** **I** **G** **Q** **G** **SD** **V** **I** **I** **G** **Q** **G** **E** **V** **D** **T** **T** **E** **D** **T** **Q** **S** **G** **M** **T** **G** **H** **S**
 GI-28810263 593 **SD** **V** **I** **I** **G** **Q** **G** **E** **V** **D** **T** **T** **E** **D** **T** **Q** **S** **G** **M** **T** **G** **H** **S** **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S** **SD** **V** **I** **I** **G** **Q** **G** **E** **V** **D** **T** **T** **E** **D** **T** **Q** **S** **G** **M** **T**
 GI-21909640 479 **SD** **V** **I** **I** **G** **Q** **G** **E** **V** **D** **T** **T** **E** **D** **T** **Q** **S** **G** **M** **T** **G** **H** **S** **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S** **SD** **V** **I** **I** **G** **Q** **G** **E** **V** **D** **T** **T** **E** **D** **T** **Q** **S** **G** **M** **T**
 GI-19224141 1021 **SD** **L** **I** **I** **G** **Q** **G** **E** **V** **D** **T** **T** **E** **D** **T** **Q** **S** **G** **M** **T** **G** **H** **S**

GI-19745307 602 --- **Q** **I** **V** **E** **T** **T** **E** **D** **T** **Q** **T** **G** **M** **H** **G** **D** **S** **G** **C** **K** **T** **E** **V** **E** **D** **T** **K** **L** **V** **Q** **S** **F** **H** **D** **N** **K**
 ORF84 602 --- **Q** **I** **V** **E** **T** **T** **E** **D** **T** **Q** **T** **G** **M** **H** **G** **D** **S** **G** **C** **K** **T** **E** **V** **E** **D** **T** **K** **L** **V** **Q** **S** **F** **H** **D** **N** **K**
 GI-28810263 619 --- **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S** **SD** **V** **I** **I** **G** **Q** **G** **Q** **I** **V** **E** **T** **T** **E** **D** **T** **Q** **T** **G** **M** **H** **G** **D** **S** **G** **C** **K** **T** **E** **V** **E** **D** **T** **K** **L** **V** **Q** **S** **F** **H** **D** **N** **K**
 GI-21909640 539 **G** **H** **S** **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S** **SD** **V** **I** **I** **G** **Q** **G** **Q** **I** **V** **E** **T** **T** **E** **D** **T** **Q** **T** **G** **M** **H** **G** **D** **S** **G** **C** **K** **T** **E** **V** **E** **D** **T** **K** **L** **V** **Q** **S** **F** **H** **D** **N** **K**
 GI-19224141 1047 --- **G** **S** **T** **T** **E** **I** **E** **D** **S** **K** **S** **SD** **V** **I** **I** **G** **Q** **G** **Q** **I** **V** **E** **T** **T** **E** **D** **T** **Q** **T** **G** **M** **H** **G** **D** **S** **G** **C** **K** **T** **E** **V** **E** **D** **T** **K** **L** **V** **Q** **S** **F** **H** **D** **N** **K**

GI-19745307 639 **E** **S** **E** **S** **N** **S** **E** **I** **P** **K** **R** **D** **K** **P** **R** **S** **N** **T** **S** **L** **P** **A** **T** **G** **E** **R** **Q** **H** **N** **M** **F** **F** **I** **M** **V** **T** **S** **C** **S** **L** **I** **S** **S** **V** **F** **V** **I** **S** **L** **K** **T** **K** **R** **L** **S** **S** **C**
 ORF84 639 **E** **S** **E** **S** **N** **S** **E** **I** **P** **K** **R** **D** **K** **P** **R** **S** **N** **T** **S** **L** **P** **A** **T** **G** **E** **R** **Q** **H** **N** **M** **F** **F** **I** **M** **V** **T** **S** **C** **S** **L** **I** **S** **S** **V** **F** **V** **I** **S** **L** **K** **T** **K** **R** **L** **S** **S** **C**
 GI-28810263 676 **E** **P** **E** **S** **N** **S** **E** **I** **P** **K** **R** **D** **K** **P** **R** **S** **N** **T** **S** **L** **P** **A** **T** **G** **E** **R** **Q** **H** **N** **K** **F** **F** **I** **M** **V** **T** **S** **C** **S** **L** **I** **S** **S** **V** **F** **V** **I** **S** **L** **K** **S** **K** **R** **L** **S** **S** **C**
 GI-21909640 599 **E** **P** **E** **S** **N** **S** **E** **I** **P** **K** **R** **D** **K** **P** **R** **S** **N** **T** **S** **L** **P** **A** **T** **G** **E** **R** **Q** **H** **N** **K** **F** **F** **I** **M** **V** **T** **S** **C** **S** **L** **I** **S** **S** **V** **F** **V** **I** **S** **L** **K** **S** **K** **R** **L** **S** **S** **C**
 GI-19224141 1104 **E** **P** **E** **S** **N** **S** **E** **I** **P** **K** **R** **D** **K** **P** **R** **S** **N** **T** **S** **L** **P** **A** **T** **G** **E** **R** **Q** **H** **N** **M** **F** **F** **I** **M** **V** **T** **S** **C** **S** **L** **I** **S** **S** **V** **F** **V** **I** **S** **L** **K** **S** **K** **R** **L** **S** **S** **C**

FIGURE 54A

GI-19224WO 2006/078318-MRKNKLLLATAILATALGTASLNQNVKAEAGVSSGOLTKRSITPCT/US2005/027239

ORF80 PCT1-LEBKE-MRKNKLLLATAILATALGTASLNQNVKAEAGVVTCKSLQVTKMT-EDDEVLN
 GI-21909636 1 -MRKNKLLLATAILATALGTASLNQNVKAEAGVSENKLVKNTFDSYTDNEVLM
 GI-28810259 1 MEREMRKNKLLLATAILATALGTASLNQNVKAEAGVSENKLVKNTFDSYTDNEVLM
 GI-19745303 1 -----MRKNKLLLATAILATALGTASLNQNVKAEAGVLDGSILVKKTFPSYTDNEVLM
 GI-13621428 1 -----MLRLHLLTCAALTSFAATT-----VHGETVNGAKLTVTNLDLVNSN-ALH

GI-19224137 56 PNTDYTFVNFDSAPTCTESN-LEIKEGIAVN-NQIKVSYSTNDKTSGNKQVVDVEMK
 ORF80 60 PETAFETFTLEPDMTASCHEGS-LDIRNGIVBGLDKQVTVKYKNTDKTSQRTNLAQFDFSN
 GI-21909636 56 PRADYTFKVEADSTASQKTKDGLKIKPGIVNGLT-EQILSYTNTDNPDSNVKSTEFDFSK
 GI-28810259 61 PRADYTFKVEADSTASQKTKDGLKIKPGIVNGLT-EQILSYTNTDNPDSNVKSTEFDFSK
 GI-19745303 56 PRADYTFKVEADSTASQKTKDGLDIRPGVLDGLKENTKTIHYCNSDKTTANERSVNFDFAN
 GI-13621428 48 PNTDYTFKVEADSTVNFDSAPTCTESN-LEIKEGIAVN-NQIKVSYSTNDKTSGNKQVVDVEMK

GI-19224137 114 VTFPSVGYRYRVVTEENKCTAE-CITYDDTKNLTVDVYVGN-NEKGGLEPKYIVSKKEDSA
 ORF80 119 VKFPAICGVYRYVMVSEKNDKED-CITYDDTKNLTVDVYVGNKANNEEGFEVLVIVSKEGESS
 GI-21909636 115 VVFPGIGVYRYTVSEKQGDVE-CITYDDTKNLTVDVYVGN-KEGGGFEPKIVSKKEDSA
 GI-28810259 120 VVFPGIGVYRYTVSEKQGDVE-CITYDDTKNLTVDVYVGN-KEGGGFEPKIVSKKEDSA
 GI-19745303 116 VKFPAICGVYRYTVSEVNCNFA-CITYDDTKNLTVDVYVGN-KEGGGFEPKIVSKKEDSA
 GI-13621428 103 VTFEKEGVYRYRYVTEENKCTAE-CITYDDTKNLTVDVYVGN-NEKGGLEPKYIVSKKEDSA

GI-19224137 171 TRPPIQFNNFETTSKTEKEVTGNTGDRKRAFTFTLTLOFNEYMEASSVVKIEENGQ--
 ORF80 178 TKKPIEFNTSLIKTTSLKTEKQITGNACDRKSENFLLTLOFSEYNTGSSVVKIEQDGS--
 GI-21909636 172 VKKPIQFNNFETTSKTKVKNVSGNTGELQKEEDFTLTLESTNEKNDQIVSLQKQNE--
 GI-28810259 177 VKKPIQFNNFETTSKTKVKNVSGNTGELQKEEDFTLTLESTNEKNDQIVSLQKQNE--
 GI-19745303 173 DKKPIQFNNFETTSKTKVKNVSGNTGELQKEEDFTLTLESTNEKNDQIVSLQKQNE--
 GI-13621428 161 -KVPPIQFNNFETTSKTKVKNVSGNTGELQKEEDFTLTLESTNEKNDQIVSLQKQNE--

GI-19224137 229 ----TRDVKIGEAVKFTLNDSSVLSKLPVGINVYKVEEAEANQGGYTTATLNDG--EK
 ORF80 236 ----KDVVIGTPYKFTLCHCKSVMLSKLPVGINVLSSEDEANKDGYTTATLNDG--EK
 GI-21909636 230 ----KEVRICTPYKFTLCHCKSVMLSKLPVGINVLSSEDEANKDGYTTATLNDG--EK
 GI-28810259 235 ----KEVRICTPYKFTLCHCKSVMLSKLPVGINVLSSEDEANKDGYTTATLNDG--EK
 GI-19745303 231 ----TRDVKIGEAVKFTLNDSSVLSKLPVGINVYKVEEAEANQGGYTTATLNDG--EK
 GI-13621428 220 QAPVTEASIDQVHTFTLNDGESIKVTNLFVGVDAVVTEDDYKSEKVTNVEVSPQDGA

GI-19224137 283 LSTVNLG-QEHTKTKTADIVVTNNRDTQVPTGVVGTLPFAVLSIVAIGGVYITRKK
 ORF80 292 SSETLSTONOKTDESADIVVTNNRDTQVPTGVVGTLPFAVLSIVAIGGVYITRKK
 GI-21909636 285 SKVOLD-MEOKTDESADIVVTNNRDTQVPTGVVGTLPFAVLSIVAIGGVYITRKK
 GI-28810259 290 SKVOLD-MEOKTDESADIVVTNNRDTQVPTGVVGTLPFAVLSIVAIGGVYITRKK
 GI-19745303 285 LSTVNLG-QEHTKTKTADIVVTNNRDTQVPTGVVGTLPFAVLSIVAIGGVYITRKK
 GI-13621428 280 KNTAGNSTEGETSDKDDMTTFTNNKDEVEPTGVVGTLPFAVLSIVAIGGVYITRKK

GI-19224137 342 A
 ORF80 352 A
 GI-21909636 344 A
 GI-28810259 349 A
 GI-19745303 344 A
 GI-13621428 340 A

FIGURE 55

173/487

GI-2190WO 2006/078318

PCT/US2005/027239

GI-28810261	1	MLFSVVMILTMLAFNOTVLAKDSTV
GI-19224139	1	MLFSVVMILTMLAFNOTVLAKDSTV
ORF82	1	MLFSVVMILTMLAFNOTVLAKDSTV
GI-19745305	1	MLFSVVMILTMLAFNOTVLAKDSTV

GI-21909638	20	QTSISVENVLERAGDSTPFSIALESIDAMNTIEBITIAGSGKASFSPLTFTTVGQYTYRV
GI-28810261	26	QTSISVENVLERAGDSTPFSIALESIDAMNTIEBITIAGSGKASFSPLTFTTVGQYTYRV
GI-19224139	26	QTSISVENVLERAGDSTPFSIALESIDAMNTIEBITIAGSGKASFSPLTFTTVGQYTYRV
ORF82	61	QTSISVENVLERAGDSTPFSIALESIDAMNTIEBITIAGSGKASFSPLTFTTVGQYTYRV
GI-19745305	32	QTSISVENVLERAGDSTPFSIALESIDAMNTIEBITIAGSGKASFSPLTFTTVGQYTYRV

GI-21909638	80	YOKPSQNKDYQADTTVFDVLVYVYTYDEDGTLVAKVISRRAGDEEKSAITFKPKRLVKPIP
GI-28810261	86	YOKPSQNKDYQADTTVFDVLVYVYTYDEDGTLVAKVISRRAGDEEKSAITFKPKRLVKPIP
GI-19224139	86	YOKPSQNKDYQADTTVFDVLVYVYTYDEDGTLVAKVISRRAGDEEKSAITFKPKRLVKPIP
ORF82	121	YOKPSQNKDYQADTTVFDVLVYVYTYDEDGTLVAKVISRRAGDEEKSAITFKPKRLVKPIP
GI-19745305	92	YOKPSQNKDYQADTTVFDVLVYVYTYDEDGTLVAKVISRRAGDEEKSAITFKPKRLVKPIP

GI-21909638	140	PRQPNIPKTPPLPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL
GI-28810261	146	PRQPNIPKTPPLPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL
GI-19224139	146	PRQPNIPKTPPLPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL
ORF82	181	PRQPNIPKTPPLPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL
GI-19745305	152	PRQPNIPKTPPLPLAGEVKSLLGILSIVLLGLLVLLYVKKLKSRL

FIGURE 56

PCT/US05/27239 174/487

!!SEQUENCE_LIST 1.0

(Peptide) FASTA of: gi-50913505.pep from: 1 to: 1036 September 15, 2004 18:46

gi|50913505|ref|YP_059477.1| Collagen adhesion protein [Streptococcus pyogenes M GAS10394]

TO: *.pep Sequences: 56 Symbols: 22,803 Word Size: 2

Scoring matrix: GenRunData:blosum50.cmp

Variable pamfactor used

Gap creation penalty: 12 Gap extension penalty: 2

Histogram Key:

Each histogram symbol represents 1 search set sequences
z-scores computed from opt scores

z-score	obs	exp
(=)	(*)	
< 20	0	0:
22	0	0:
24	0	0:
26	0	0:
28	0	0:
30	0	0:
32	0	0:
34	0	1:*
36	0	2: *
38	0	3: *
40	0	4: *
42	0	5: *
44	0	5: *
46	5	5:====*
48	10	5:====*=====
50	12	4:====*=====
52	4	4:====*
54	6	3:====*
56	4	3:====*
58	5	2:====*
60	3	2:====*
62	0	1:*
64	1	1:*
66	1	1:*
68	0	1:*
70	1	1:*
72	0	0:
74	0	0:
76	1	0:==
78	0	0:
80	0	0:
82	0	0:
84	0	0:
86	0	0:
88	0	0:
90	0	0:
92	0	0:
94	0	0:
96	1	0:==
98	0	0:

FIGURE 57

PCT/US05/27239

```

100 0 0:
102 0 0:
104 0 0:
106 1 0:=
108 0 0:
110 0 0:
112 0 0:
114 0 0:
116 0 0:
118 0 0:
>120 1 0:=

```

Joining threshold: 39, opt. threshold: 27, opt. width: 16, reg.-scaled

The best scores are:

initl initn opt z-sc E(54)...

```

/home/morاما/gas/pili/align/gi-50913505.pep Begin: 1 End: 1036
! gi|50913505|ref|YP_059477.1| Collag... 6697 6697 6697 3452.1 9.6e-189
/home/morاما/gas/pili/align/gi-19224141.pep Begin: 48 End: 144
! gi|19224141|gb|AAL86412.1|AF447492_... 63 100 159 105.9 0.023
/home/morاما/gas/pili/align/gi-21909640.pep Begin: 147 End: 449
! gi|21909640|ref|NP_663908.1| protei... 35 35 136 96.3 0.08
/home/morاما/gas/pili/align/gi-13621428.pep Begin: 57 End: 318
! gi|13621428|gb|AAK33238.1| hypothet... 33 33 91 75.6 1.1
/home/morاما/gas/pili/align/gi-50913506.pep Begin: 33 End: 428
! gi|50913506|ref|YP_059478.1| Fimbri... 70 149 86 71.3 1.9
/home/morاما/gas/pili/align/gi-13621432.pep Begin: 14 End: 56
! gi|13621432|gb|AAK33241.1| conserve... 40 65 78 68.0 2.9
/home/morاما/gas/pili/align/gi-19745301.pep Begin: 241 End: 466
! gi|19745301|ref|NP_606437.1| putati... 52 52 73 64.8 4.3
/home/morاما/gas/pili/align/gas15.pep Begin: 492 End: 739
! GAS15 GAS15 43 68 69 61.4 6.6
/home/morاما/gas/pili/align/gi-21909636.pep Begin: 176 End: 298
! gi|21909636|ref|NP_663904.1| conser... 31 31 62 60.8 7.1
/home/morاما/gas/pili/align/gi-28810259.pep Begin: 181 End: 303
! gi|28810259|dbj|BAC63197.1| hypothe... 31 31 62 60.7 7.2
/home/morاما/gas/pili/align/gi-19224139.pep Begin: 90 End: 143
! gi|19224139|gb|AAL86410.1|AF447492_... 43 43 54 58.9 8.9
/home/morاما/gas/pili/align/gi-19745305.pep Begin: 96 End: 149
! gi|19745305|ref|NP_606441.1| hypoth... 43 43 54 58.8 9
/home/morاما/gas/pili/align/orf82.pep Begin: 125 End: 178
! TRANSLATE of: orf82.seq check: 4296... 43 43 54 58.2 9.6
/home/morاما/gas/pili/align/gi-21909638.pep Begin: 84 End: 137
! gi|21909638|ref|NP_663906.1| hypoth... 43 43 52 58.0 9.9
\\End of List

```

gi-50913505.pep

/home/morاما/gas/pili/align/gi-50913505.pep

gi|50913505|ref|YP_059477.1| Collagen adhesion protein [Streptococcus pyogenes M GAS10394]

SCORES Initl: 6697 Initn: 6697 Opt: 6697 z-score: 3452.1 E(): 9.6e-189
 >>/home/morاما/gas/pili/align/gi-50913505.pep (1036 aa)
 initn: 6697 initl: 6697 opt: 6697 Z-score: 3452.1 expect(): 9.6e-189
 Smith-Waterman score: 6697; 100.0% identity in 1036 aa overlap
 (1-1036:1-1036)

10 20 30 40 50 60
 gi-50913505. MYSRLKRELIVIVINRKKKYKLIRLMVTVGLIFSQVLVPIRRLGLQMISTQTKVIPQEIIVT

FIGURE 57A

FIGURE 57B

FIGURE 57C.

FIGURE 57D

FIGURE 57E

initn: 149 initl: 70 opt: 86 Z-score: 71.3 expect(): 1.9
 Smith-Waterman score: 120; 21.5% identity in 469 aa overlap
 (503-966:33-428)

```

      480      490      500      510      520
gi-50913505. YSQWVYHGTFDYSSYQGESFNRGQIHRYRGIVSVSDGIRDDAVKNSLLGVNG-----L
      |:: || :: :||::: :: :::: |
gi-50913506. NRRETVREKILITAKKLMACILAVVGLGMRVS-ALSKDDTAQLKITNIEGGPTVTL
      10      20      30      40      50      60.

      530      540      550      560      570      580
gi-50913505. LQRFVNINPENKLSVIGFQGSADYHAGKWYPDQSPRGGFYQPNLNNSRDAELLKGWSTNS
      : :: | | |:: | | : : | : : | : : | : |::
gi-50913506. YKIGEGVYNTNGDSFINFK-----YAEGVSLTETGPTSQEIT-TIANGINTGKIKPFSTEN
      70      80      90      100     110

      590      600      610      620      630      640
gi-50913505. LLDPNLTALHNNGTNYHAALLKAKEILNEVKDDGRRKIMIFISDGVPTFYFGEDGYRSG
      : | :: : |:: : ||| : ||| |:: : | || : :
gi-50913506. VSISNGTATYNARGASVYIALLTGAT-----DGRTYNPILLAAS-----YNGEGNLVTK
      120     130     140     150     160

      650      660      670      680      690      700
gi-50913505. NGSSNDRNNVTRSQEGSKLAIDEFKARYPNLSIYSLGVSKDINSDTASSSPVVLKYLSGE
      | : | :: | : | :: | |; |:: : | : |:: | : | : | : |
gi-50913506. NIDS--KSNLYGQ--TSVA---KSSLPSITKKVTGTIDDVNKKTTSLGSVLSYSLTFE
      170     180     190     200     210

      710      720      730      740      750      760
gi-50913505. EHYYGITDTABLEKTLNKIVEDSKLSQLGISDSLSQYVDYDKQPDVLVTRKSKVNDETE
      | |::: || | : ||::: : : : : : | |::: ||
gi-50913506. LPSY-----TKEAVNKTVY-----VSDNMSEGLTF--NFNSLTVEWKGMANITE
      220     230     240     250     260

      770      780      790      800      810      820
gi-50913505. ILYQKDQVQEAGKDIIDKVFTPKTTSQPKGKVTLTFFKSDYKVDDEYTYTILSFNVKASDE
      | |::: | : | : | : | : | : | :
gi-50913506. -----DGSVMVENTKIGIAKEVNNGFNLSFIYDS--LESISPNI-----
      270     280     290

      830      840      850      860      870      880
gi-50913505. AYEKYKDNEGRYSEMCDSDTDYGTNQTSSGKGGLPSNSDASVNYMADGREQKLPYKHPVI
      : | : |:: : | | : : : | | : | : | : | : |
gi-50913506. SYKAVVNKKAIVCEEGNPNKAEFFYSNNPTKGNLYDNLDKKPKK-CNGITSKEDSK---
      300     310     320     330     340     350

      890      900      910      920      930      940
gi-50913505. QVKTVPITFTTKVDADNNQKLAGVEFELRKEDKKIWEKGTGSGNQLNFKYLQKGKTY
      | | | : | ||: : : | | : : | : | : | : |
gi-50913506. IVITYQIAERKVDV-SKTPILGALFGVYDTSNKLI-DIVTINKNGYAISTQVSSGK-YK
      360     370     380     390     400

      950      960      970      980      990     1000
gi-50913505. LYETKAKLGYTLPENPWEVAVANNGDIKVKHPIEGELKSKDGSYMIKNYKIYQLPSSGGR
      : | | | | : | : : | :
gi-50913506. IKELKAPKGYSLNTETYEITANWVTATVKTSANSKSTTYTSDKNKATDNSEQVGWLKNGI
      410     420     430     440     450     460

```

gi-50913505.pep

/home/morama/gas/pili/align/gi-13621432.pep

gi|13621432|gb|AAK33241.1| conserved hypothetical protein [Streptococcus pyogene

FIGURE 57F

PCT/US2005/027239/181/487

s]

SCORES Init1: 40 Initn: 65 Opt: 78 z-score: 68.0 E(): 2.9
 >>/home/morama/gas/pili/align/gi-13621432.pep (450 aa)
 initn: 65 init1: 40 opt: 78 z-score: 68.0 expect(): 2.9
 Smith-Waterman score: 78; 37.0% identity in 46 aa overlap
 (368-411:14-56)

	340	350	360	370	380	390
gi-50913505.	KGHNNTINIYIRSTRPLGLKPIGAAPALIQPRSFRLTPRSTRMK--RSAPVEKFEGELE					
					::	:: : : :
gi-13621432.	MTRTNYQKKRMTCPVETEDITYRRKKIKGRRQAILAQFEPELV					
				10	20	30 40

	400	410	420	430	440	450
gi-50913505.	HHKRIDYLGDNQNNPDITIDDKEDHDTSDLYRLYLDMTGKKNPLDIEVVVDKSGSMQEG					
	:	: :				
gi-13621432.	HHELI---GDSCPCDCHGTLTEIGSVVQRQELVFIPAQLKRINHVVQHAYKCQTCSDNSL					
	50	60	70	80	90	100

gi-50913505.pep

/home/morama/gas/pili/align/gi-19745301.pep

gi|19745301|ref|NP_606437.1| putative collagen binding protein [Streptococcus pyogenes MGAS8232]

SCORES Init1: 52 Initn: 52 Opt: 73 z-score: 64.8 E(): 4.3
 >>/home/morama/gas/pili/align/gi-19745301.pep (524 aa)
 initn: 52 init1: 52 opt: 73 z-score: 64.8 expect(): 4.3
 Smith-Waterman score: 95; 23.7% identity in 245 aa overlap
 (759-989:241-466)

	730	740	750	760	770	780
gi-50913505.	DSKLSQLGISDSLSQYVDYYDKQPDVLVTRKSKVNDETEILYQKDQVQVQAG-KDIIDKVV					
					:::	: : : :::
gi-19745301.	ETIDPDFNEGKEIKYTHILGADLFSYANNPRASTNDE--LLSQVKKVLEKGYRD--DSTT					
	220	230	240	250	260	

	790	800	810	820	830	840
gi-50913505.	FTPKTTSQPKGKVTLT--FKSDYKVD--EYTYLSEFNVKASDEAYEKYKDNENGRYSEM					
	:::	:	:	:	:	:
gi-19745301.	YANLTSVEFRAATQLAIYYFTDSVDLDNLADYHGFALTEALNATKEIVAYAEADRANLP					
	270	280	290	300	310	320

	850	860	870	880	890
gi-50913505.	GDSDTDY---GTNQTSS--GKGGLPSNSDASVNYMADGREQKLPHYKHPVIQVKTIVPITFT				
	:	:	:	:	:
gi-19745301.	NISNLDFYVPNSNKYQSLIGTQYHP--ESLVDIIRMEDKQAPIIPITHKLTISKITVTGTI-				
	330	340	350	360	370 380

	900	910	920	930	940	950
gi-50913505.	KVDADNNQKKLAGVEFELRKEDKKIVWEKGTGSGN-GQINFKYLQKQK-TYYLYETKAKL					
		:	:	:	:	:
gi-19745301.	---AD--KKKEFNFEIHLKSSDQAI--SGTYPTNSGELT---VTDGKATFTLKDGESLI					
	390	400	410	420	430	

	960	970	980	990	1000	1010
gi-50913505.	GYTLPEN-PWEVAVANNGDIKVKHPIEGELKSKDGSYMIKNYKIYQLPSSGGRGSQIFII					
		:	:	:	:	:
gi-19745301.	VEGLPSGYSYEITETGASDYEVS--VNGK-NAPDGKATKASVKEDETITFENRKDLVPPT					

FIGURE 57G

PCT/US05/27239

440 450 460 470 480 490

1020 1030

gi-50913505. VGSMTATVALLFYRRQHRKKQY

gi-19745301. GLTTDGAIIYLWLLLLVLLGLWVWLIGRKLKND

500 510 520

gi-50913505.pep

/home/morama/gas/pili/align/gas15.pep

GAS15 GAS15

SCORES Init1: 43 Initn: 68 Opt: 69 z-score: 61.4 E(): 6.6

>>/home/morama/gas/pili/align/gas15.pep (762 aa)

initn: 68 init1: 43 opt: 69 Z-score: 61.4 expect(): 6.6

Smith-Waterman score: 100; 21.4% identity in 252 aa overlap

(641-873:492-739)

```

          620      630      640      650      660
gi-50913505. AKEILNEVKDDGRRKIMIFISDGVPTFYFGEDGYRSGNGSSNDRNNVTRSQ--EGSKLAI
              |||| :|:: : ::|::| :::|||
gas15.pep     HIAGRDLFKYTVKPRDTPDPTFLKHIIKKVIEKGYRE-KGQAI EYSGLTETQLRAATQLAI
              470      480      490      500      510      520

```

```

          670      680      690      700      710      720
gi-50913505. DEF--KARYPNLSIYSLGVSKDINSSTASSPVVLKYLSGEEHYGYITDTAELEKTLNKI
              |:: : : : : |::| : : : : : |:: : : : : : |:: : : : : : |
gas15.pep     YYFTDSAE LDKDKLDYHGFGDMNDSTLAVAKILVEY-AQDSNPPQLTDLDFIPNNNKY
              530      540      550      560      570

```

```

          730      740      750      760      770
gi-50913505. VEDSKLSQLGISDSLSQYVDYYDKQPDVLVT----RKSQVN---DETEILYQKDQVQEA
              :| : : : : : ||: : || : : : : ||: : : : : : |:: : : : : :
gas15.pep     --QSLIGTQWHPEDLVDIIRMEDKKEVIPVTHNLTLRKTVTGLAGDRTKDFHFEIELKNN
              580      590      600      610      620      630

```

```

          780      790      800      810      820      830
gi-50913505. GKDIIDKVVFVTPKTTSQPK-GKVTILTFKSDYKVDDE-YTYTILSFNVKASDEAYEKYKDNE
              : : : : : | | | : : | | : : : : : : : : : : | : | : | |
gas15.pep     KQELLSQTVKTDKTNLEFKDGKATINLKHGESLTLOGLPEGYSYLVKETDSEGYKVKVNS
              640      650      660      670      680      690

```

```

          840      850      860      870      880      890
gi-50913505. GRYSMGDSDDTYGTNOT----SSGKGLPSNSDASVN-YMADGREOKLPYKHPVIQVKT
              : : : | | : : | : : : : : : | : : | : |
gas15.pep     QEVANATVSKTGITSDETAFENNKPEVVPVTVGVQKINGYLALIVIAGISLGIWGIHTIR
              700      710      720      730      740      750

```

```

          900      910      920      930      940      950
gi-50913505. VPITFTKVDADNNQKKLAGVEFELRKEDKKIVWEKGTGSGNQLNFKYLQKGTYYLYET
gas15.pep     IRKHD
              760

```

gi-50913505.pep

/home/morama/gas/pili/align/gi-21909636.pep

gi|21909636|ref|NP_663904.1| conserved hypothetical protein [Streptococcus pyogenes MGAS315]

FIGURE 57H

183/487

SCORES Init1: 31 Initn: 31 Opt: 62 z-score: 60.8 E(): 7.1
 >>/home/morama/gas/pili/align/gi-21909636.pep (344 aa)
 initn: 31 init1: 31 opt: 62 Z-score: 60.8 expect(): 7.1
 Smith-Waterman score: 71; 22.9% identity in 131 aa overlap
 (181-305:176-298)

```

      160      170      180      190      200      210
gi-50913505. DITTTTLTFKNGLSLEGASTEANDPNVRVGI VNPNDTVQTITPTIKQDADGKVNILVFTGR
      || |:: | : :|:::|: ||
gi-21909636. VDVYVGNKEGGGFEPKFIVSKEQGT DVKKPVNFNNSFATTSLKVKKNVSGN-----TGE
      150      160      170      180      190

      220      230      240      250      260
gi-50913505. LGKQ----VIIVSTTRLKEEQTISLDSYGELVIDGAVGLSQKDRPPYSKPITVNLKPKL
      | | : : : | :|::| :||:: | : : : | : : : | : : | : : |
gi-21909636. LQKEFDFTLTINESTNFKKDQIVSLQK-GNEKFEVKIGTPYKFKLNKGESIQLDKLPVGI
      200      210      220      230      240      250

      270      280      290      300      310      320
gi-50913505. SSISSLDSKDFEIVKTI DNLYTWDDQ--FYLLDFISKQYEV LKTDYQSAKDSTPQTRDI
      : : : : : || :| | | :| ||: | |
gi-21909636. TYKVNEMEANK-DGYKTTASLKEGDGQSKMYQLDMEQKTDESAD EIVVTNKRDTQVPTGV
      260      270      280      290      300      310

      330      340      350      360      370      380
gi-50913505. LFGEYTV EPLVMNKGHNNTINIYIRSTRLGLKPIGAAPALIQPRSF RSLTPRSTRMKRS
gi-21909636. VGT LAPFAVLSIVAIGGVIIYITKRKKA
      320      330      340

```

gi-50913505.pep
 /home/morama/gas/pili/align/gi-28810259.pep

gi|28810259|dbj|BAC63197.1| hypothetical protein [Streptococcus pyogenes SSI-1]

SCORES Init1: 31 Initn: 31 Opt: 62 z-score: 60.7 E(): 7.2
 >>/home/morama/gas/pili/align/gi-28810259.pep (349 aa)
 initn: 31 init1: 31 opt: 62 Z-score: 60.7 expect(): 7.2
 Smith-Waterman score: 71; 22.9% identity in 131 aa overlap
 (181-305:181-303)

```

      160      170      180      190      200      210
gi-50913505. DITTTTLTFKNGLSLEGASTEANDPNVRVGI VNPNDTVQTITPTIKQDADGKVNILVFTGR
      || |:: | : :|:::|: ||
gi-28810259. VDVYVGNKEGGGFEPKFIVSKEQGT DVKKPVNFNNSFATTSLKVKKNVSGN-----TGE
      160      170      180      190      200

      220      230      240      250      260
gi-50913505. LGKQ----VIIVSTTRLKEEQTISLDSYGELVIDGAVGLSQKDRPPYSKPITVNLKPKL
      | | : : : | :|::| :||:: | : : : | : : : | : : | : : |
gi-28810259. LQKEFDFTLTINESTNFKKDQIVSLQK-GNEKFEVKIGTPYKFKLNKGESIQLDKLPVGI
      210      220      230      240      250      260

      270      280      290      300      310      320
gi-50913505. SSISSLDSKDFEIVKTI DNLYTWDDQ--FYLLDFISKQYEV LKTDYQSAKDSTPQTRDI
      : : : : : || :| | | :| ||: | |
gi-28810259. TYKVNEMEANK-DGYKTTASLKEGDGQSKMYQLDMEQKTDESAD EIVVTNKRDTQVPTGV
      270      280      290      300      310      320

```

FIGURE 57I

PCT/US05/27239

330 340 350 360 370 380
 gi-50913505. LFGEYTVPEPLVMNKGHNNTINIIYIRSTRPIGLKPIGAAPALIQPRSFRLTPRSTRMKRS

gi-28810259. VGTLPAPFAVLSIVAIGGVIIYITKRKKA
 330 340

gi-50913505.pep
 /home/morama/gas/pili/align/gi-19224139.pep

gi|19224139|gb|AAL86410.1|AF447492_7 unknown [Streptococcus pyogenes]

SCORES Init1: 43 Initn: 43 Opt: 54 z-score: 58.9 E(): 8.9
 >>/home/morama/gas/pili/align/gi-19224139.pep (189 aa)
 initn: 43 init1: 43 opt: 54 Z-score: 58.9 expect(): 8.9
 Smith-Waterman score: 54; 31.6% identity in 57 aa overlap
 (742-796:90-143)

	720	730	740	750	760
gi-50913505.	GITDTAELEKTLNKIVEDSKLSQLGISDSLSQYVDYYDKQP--DVLVTRKSKVNDETEIL				
					:
gi-19224139.	ITIAGSGKASFSPITFTTVGQYTYRVYQKPSQNKDYQADTTVFDVLVYV---TYDEDGTL				
	60	70	80	90	100 110

	770	780	790	800	810	820
gi-50913505.	YQKDQVQEAGKDIIDKVVFPTKTTSPQKGVTLTFKSDYKVDDEYTYTSLFNVKASDEAY					
		:: :	: ::	:		
gi-19224139.	VAKVISRRAGDEEKSAITFKPKRLVKPIPPROPNIPKTPPLAGEVKSLLGILSIVLLGL					
	120	130	140	150	160	170

gi-50913505.pep
 /home/morama/gas/pili/align/gi-19745305.pep

gi|19745305|ref|NP_606441.1| hypothetical protein [Streptococcus pyogenes MGAS82 32]

SCORES Init1: 43 Initn: 43 Opt: 54 z-score: 58.8 E(): 9
 >>/home/morama/gas/pili/align/gi-19745305.pep (195 aa)
 initn: 43 init1: 43 opt: 54 Z-score: 58.8 expect(): 9
 Smith-Waterman score: 54; 31.6% identity in 57 aa overlap
 (742-796:96-149)

	720	730	740	750	760
gi-50913505.	GITDTAELEKTLNKIVEDSKLSQLGISDSLSQYVDYYDKQP--DVLVTRKSKVNDETEIL				
					:
gi-19745305.	ITIAGSGKASFSPITFTTVGQYTYRVYQKPSQNKDYQADTTVFDVLVYV---TYDEDGTL				
	70	80	90	100	110 120

	770	780	790	800	810	820
gi-50913505.	YQKDQVQEAGKDIIDKVVFPTKTTSPQKGVTLTFKSDYKVDDEYTYTSLFNVKASDEAY					
		:: :	: ::	:		
gi-19745305.	VAKVISRRAGDEEKSAITFKPKRLVKPIPPROPDIPKTPPLAGEVKSLLGILSIVLLGL					
	130	140	150	160	170	180

gi-50913505.pep
 /home/morama/gas/pili/align/orf82.pep

TRANSLATE of: orf82.seq check: 4296 from: 1 to: 672
 generated symbols 1 to: 224.

FIGURE 57J

PCT/US2005/027239
GETSEQ from morama, September 13, 2004 17:09.

SCORES Init1: 43 Initn: 43 Opt: 54 z-score: 58.2 E(): 9.6
>>/home/morama/gas/pili/align/orf82.pep (224 aa)
initn: 43 init1: 43 opt: 54 Z-score: 58.2 expect(): 9.6
Smith-Waterman score: 54; 31.6% identity in 57 aa overlap
(742-796:125-178)

	720	730	740	750	760
gi-50913505.	GITDTAELEKTLNKIVEDSKLSQLGISDSLSQYVDYYDKQP--DVLVTRKSKVNDETEIL				
					:
orf82.pep	ITIAGSGKASFSPLTFTTVGQYTYRVYQKPSQNKDYQADTTVFDVLVYV---TYDEDGTL				
	100	110	120	130	140 150

	770	780	790	800	810	820
gi-50913505.	YQKDQVQEAGKDIIDKVVFTPKTTSQPKGKVTLTFSKDYKVDDEYTYTLSFNVKASDEAY					
		:: :	: :: :			
orf82.pep	VAKVISRRAGDEEKSAITFKPKRLVKPIPPRQPNIPKTPLPLAGEVKSLLGILSIVLLGL					
	160	170	180	190	200	210

gi-50913505.pep
/home/morama/gas/pili/align/gi-21909638.pep

gi|21909638|ref|NP_663906.1| hypothetical protein [Streptococcus pyogenes MGAS315]

SCORES Init1: 43 Initn: 43 Opt: 52 z-score: 58.0 E(): 9.9
>>/home/morama/gas/pili/align/gi-21909638.pep (183 aa)
initn: 43 init1: 43 opt: 52 Z-score: 58.0 expect(): 9.9
Smith-Waterman score: 52; 31.6% identity in 57 aa overlap
(742-796:84-137)

	720	730	740	750	760
gi-50913505.	GITDTAELEKTLNKIVEDSKLSQLGISDSLSQYVDYYDKQP--DVLVTRKSKVNDETEIL				
					:
gi-21909638.	ITIAGSGKASFSPLTFTTVGQYTYRVYQKPSQNKDYQADTTVFDVLVYV---TYDEDGTL				
	60	70	80	90	100 110

	770	780	790	800	810	820
gi-50913505.	YQKDQVQEAGKDIIDKVVFTPKTTSQPKGKVTLTFSKDYKVDDEYTYTLSFNVKASDEAY					
		:: :	: :: :			
gi-21909638.	VAKVISRRAGDEEKSAITFKPKRLVKPIPPRQPNIPKTPLPLAGEVKSLLGILSIVLLGL					
	120	130	140	150	160	170

! Distributed over 1 thread.

! Start time: Wed Sep 15 18:46:18 2004

! Completion time: Wed Sep 15 18:46:21 2004

! CPU time used:

! Database scan: 0:00:00.1

! Post-scan processing: 0:00:02.2

! Total CPU time: 0:00:02.3

! Output File: gi-50913505.fasta

FIGURE 57K

PCT/US2005/027239
 !!SEQUENCE_LIST 1.0

(Peptide) FASTA of: gi-50913506.pep from: 1 to: 556 September 15, 2004 18:45

gi|50913506|ref|YP_059478.1| Fimbrial structural subunit [Streptococcus pyogenes
 MGAS10394]

TO: *.pep Sequences: 56 Symbols: 22,803 Word Size: 2

Scoring matrix: GenRunData:blosum50.cmp

Variable pamfactor used

Gap creation penalty: 12 Gap extension penalty: 2

Histogram Key:

Each histogram symbol represents 1 search set sequences
 z-scores computed from opt scores

z-score	obs	exp
(=)	(*)	
< 20	0	0:
22	0	0:
24	0	0:
26	0	0:
28	0	0:
30	0	0:
32	0	0:
34	0	1:*
36	0	2: *
38	0	3: *
40	0	4: *
42	0	5: *
44	0	5: *
46	1	5:= *
48	8	5:=====*
50	24	5:=====*
52	2	4:== *
54	3	3:==*
56	1	3:= *
58	2	2:==*
60	3	2:==*
62	0	2: *
64	2	1:*=
66	2	1:*=
68	1	1:*
70	2	1:*=
72	0	0:
74	3	0:==
76	1	0:=
78	0	0:
80	0	0:
82	0	0:
84	0	0:
86	0	0:
88	0	0:
90	0	0:
92	0	0:
94	0	0:
96	0	0:
98	0	0:

FIGURE 58

```

100      0      0:
102      0      0:
104      0      0:
106      0      0:
108      0      0:
110      0      0:
112      0      0:
114      0      0:
116      0      0:
118      0      0:
>120     1      0:=

```

Joining threshold: 37, opt. threshold: 25, opt. width: 16, reg.-scaled

The best scores are:

initl initn opt z-sc E(55)...

```

/home/morana/gas/pili/align/gi-50913506.pep  Begin: 1 End: 556
! gi|50913506|ref|YP_059478.1| Fimbri... 3454 3454 3454 1016.2 4.7e-53
/home/morana/gas/pili/align/orf84.pep  Begin: 316 End: 567
! TRANSLATE of: orf84.seq check: 7868... 57 83 135 75.1 1.2
/home/morana/gas/pili/align/gi-19745307.pep  Begin: 316 End: 567
! gi|19745307|ref|NP_606443.1| protei... 57 83 135 75.1 1.2
/home/morana/gas/pili/align/gi-21909640.pep  Begin: 202 End: 524
! gi|21909640|ref|NP_663908.1| protei... 56 81 134 75.0 1.2
/home/morana/gas/pili/align/gi-28810263.pep  Begin: 316 End: 638
! gi|28810263|dbj|BAC63201.1| protein... 56 82 134 74.7 1.3
/home/morana/gas/pili/align/orf80.pep  Begin: 49 End: 352
! TRANSLATE of: orf80.seq check: 9824... 45 69 113 70.8 2.1
/home/morana/gas/pili/align/gi-19224137.pep  Begin: 25 End: 342
! gi|19224137|gb|AAL86408.1|AF447492... 45 69 109 69.8 2.4
/home/morana/gas/pili/align/gi-19224141.pep  Begin: 277 End: 645
! gi|19224141|gb|AAL86412.1|AF447492... 73 73 118 68.9 2.7
/home/morana/gas/pili/align/gi-21909636.pep  Begin: 44 End: 344
! gi|21909636|ref|NP_663904.1| conser... 45 98 96 66.1 3.8
/home/morana/gas/pili/align/gi-28810259.pep  Begin: 49 End: 349
! gi|28810259|dbj|BAC63197.1| hypothe... 45 98 96 66.0 3.8
/home/morana/gas/pili/align/gas15.pep  Begin: 222 End: 470
! GAS15 GAS15 42 68 96 63.8 5
/home/morana/gas/pili/align/gi-13621428.pep  Begin: 17 End: 340
! gi|13621428|gb|AAK33238.1| hypothet... 41 41 87 63.6 5.2
/home/morana/gas/pili/align/gi-19224135.pep  Begin: 193 End: 462
! gi|19224135|gb|AAL86406.1|AF447492... 41 41 86 61.0 7
/home/morana/gas/pili/align/gi-50913505.pep  Begin: 503 End: 966
! gi|50913505|ref|YP_059477.1| Collag... 70 149 86 60.1 7.8
/home/morana/gas/pili/align/gi-13621430.pep  Begin: 60 End: 143
! gi|13621430|gb|AAK33240.1| hypothet... 43 67 67 59.2 8.7
/home/morana/gas/pili/align/gi-19745303.pep  Begin: 44 End: 344
! gi|19745303|ref|NP_606439.1| hypoth... 51 106 69 58.4 9.5
\\End of List

```

gi-50913506.pep

/home/morana/gas/pili/align/gi-50913506.pep

gi|50913506|ref|YP_059478.1| Fimbrial structural subunit [Streptococcus pyogenes MGAS10394]

SCORES Initl: 3454 Inith: 3454 Opt: 3454 z-score: 1016.2 E(): 4.7e-53

>>/home/morana/gas/pili/align/gi-50913506.pep (556 aa)

initn: 3454 initl: 3454 opt: 3454 Z-score: 1016.2 expect(): 4.7e-53

Smith-Waterman score: 3454; 100.0% identity in 556 aa overlap

FIGURE 58A

PCT/US2005/027239

(1-556:1-556)

	10	20	30	40	50	60
gi-50913506.	MTNRRET	VREKILITAKK	MLACLAILAVV	GLGMTRVSALS	KDDTAQLKITN	IEGGPTVT
gi-50913506.	MTNRRET	VREKILITAKK	MLACLAILAVV	GLGMTRVSALS	KDDTAQLKITN	IEGGPTVT
	10	20	30	40	50	60
	70	80	90	100	110	120
gi-50913506.	LYKIGEGV	YNTNGDSFIN	KYAEGVSLTET	GPTSQEITTI	ANGINTGKIK	PFSTENV
gi-50913506.	LYKIGEGV	YNTNGDSFIN	KYAEGVSLTET	GPTSQEITTI	ANGINTGKIK	PFSTENV
	70	80	90	100	110	120
	130	140	150	160	170	180
gi-50913506.	NGTATYN	NARGASVYI	ALLTGATDGR	TYNPILLAAS	YNGEENLVTK	NIDSKSNY
gi-50913506.	NGTATYN	NARGASVYI	ALLTGATDGR	TYNPILLAAS	YNGEENLVTK	NIDSKSNY
	130	140	150	160	170	180
	190	200	210	220	230	240
gi-50913506.	AKSSSLPS	ITKKVTGTID	DVNKKTSLGS	VLSSLSLTFE	LPSTYKEAVN	KTVYVSDNM
gi-50913506.	AKSSSLPS	ITKKVTGTID	DVNKKTSLGS	VLSSLSLTFE	LPSTYKEAVN	KTVYVSDNM
	190	200	210	220	230	240
	250	260	270	280	290	300
gi-50913506.	TFNFNSLT	VEWKGKMAN	ITEDGSVMV	ENTKIGIAKE	VNNGFNLSFI	YDSLESISP
gi-50913506.	TFNFNSLT	VEWKGKMAN	ITEDGSVMV	ENTKIGIAKE	VNNGFNLSFI	YDSLESISP
	250	260	270	280	290	300
	310	320	330	340	350	360
gi-50913506.	AVVNNKA	IVGEEGNPN	KAEFFYSNN	PTKGN TYDN	LDKKPKDKN	GITSKEDSK
gi-50913506.	AVVNNKA	IVGEEGNPN	KAEFFYSNN	PTKGN TYDN	LDKKPKDKN	GITSKEDSK
	310	320	330	340	350	360
	370	380	390	400	410	420
gi-50913506.	FRKVD	SVSKTPLIGA	IFGVYDTSN	KLIDIVT	TNKNGYAIST	QVSSGKYK
gi-50913506.	FRKVD	SVSKTPLIGA	IFGVYDTSN	KLIDIVT	TNKNGYAIST	QVSSGKYK
	370	380	390	400	410	420
	430	440	450	460	470	480
gi-50913506.	LNTE	TYEITANW	VITATVKTS	ANSKSTTYT	SDKNKATDN	SEQVGWLK
gi-50913506.	LNTE	TYEITANW	VITATVKTS	ANSKSTTYT	SDKNKATDN	SEQVGWLK
	430	440	450	460	470	480
	490	500	510	520	530	540
gi-50913506.	DVKEA	YIESTKAL	TGTTFSKS	NEGSGTVL	LETDPNTKL	GELPSTGS
gi-50913506.	DVKEA	YIESTKAL	TGTTFSKS	NEGSGTVL	LETDPNTKL	GELPSTGS
	490	500	510	520	530	540
	550					
gi-50913506.	AMIGA	IGIYIVK	RRKA			
gi-50913506.	AMIGA	IGIYIVK	RRKA			
	550					

FIGURE 58B

189/487
PCT/US05/27239
gi-50913506.pep
/home/morama/gas/pili/align/orf84.pep

TRANSLATE of: orf84.seq check: 7868 from: 1 to: 2088
generated symbols 1 to: 696.
GETSEQ from morama, September 13, 2004 17:07.

SCORES Init1: 57 Initn: 83 Opt: 135 z-score: 75.1 E(): 1.2
>>/home/morama/gas/pili/align/orf84.pep (696 aa)
initn: 83 init1: 57 opt: 135 Z-score: 75.1 expect(): 1.2
Smith-Waterman score: 146; 24.4% identity in 262 aa overlap
(232-462:316-567)

gi-50913506.	210	220	230	240	250	
	KTTSLGSLVSLTFELPSYTKAEVNTVYVSDNMSEGLTFNFNSLTVEWKGKMAN----					
orf84.pep	EKEVAVDDAELKQINSEGGQEI SVTWTNQLVTDE--KGMAYIYSVKEVDKNGELLEPKDY					
	290	300	310	320	330	340
gi-50913506.	260	270	280	290	300	310
	ITEDGSVMVENTKIGIAKEVNNGFNLSFIY-----DSLESISPNISYKAVVNKAIVGEE					
orf84.pep	IKKEDGLTNTYV---KPTSGHYDIEVTFGNHIDITEDTTPDI-VSGENQMKQIEGED					
	350	360	370	380	390	
gi-50913506.	320	330	340	350		
	GNP-----NKAFFYSNNP-----TKGNTYDNL-DKKP-DKNGGITSKEDSKIIVTYQ					
orf84.pep	SKPIDEVTENNLIEFGKNTMPGEEDGTNSNKYEEVEDSRPVDTLGSLSEGGQSGDMTIE					
	400	410	420	430	440	450
gi-50913506.	360	370	380	390	400	410
	-----IAFRKVDVSKTEPLIGAIFGVYDTSNKLIDIVTTNKNGYAISTQVSSGKYIKE					
orf84.pep	EDSATHIKFSKRD-IDGKELAGATMELRDSSGKTIS--TWISDGQVKDFYLMGKYTFVE					
	460	470	480	490	500	510
gi-50913506.	420	430	440	450	460	470
	LKAPKGYSLNTE-TYEITAN-WVTATVKTSANSKSTTYTSDKNKATDNSEQVGWLKNGIF					
orf84.pep	TAAPDGYEIATAITFTVNEQGQVTVNGKATKGDAHVMV-DAYKPTKSGQVIEDIEEKLIP					
	520	530	540	550	560	570
gi-50913506.	480	490	500	510	520	530
	YSIDSRPTGNDVKEAYIESTKALTGGTFESKSNESGTFVLETDIPNTELGELPSTGSGIG					
orf84.pep	DEQCHSGSTTEIEDSKSSDVLIGGQGVIVETTEDTGTGMHGDGCKTEVEDTKIVQSFHF					
	580	590	600	610	620	630

gi-50913506.pep
/home/morama/gas/pili/align/gi-19745307.pep

gi|19745307|ref|NP_606443.1| protein F2-like protein [Streptococcus pyogenes MGA S8232]

SCORES Init1: 57 Initn: 83 Opt: 135 z-score: 75.1 E(): 1.2
>>/home/morama/gas/pili/align/gi-19745307.pep (696 aa)
initn: 83 init1: 57 opt: 135 Z-score: 75.1 expect(): 1.2
Smith-Waterman score: 146; 24.4% identity in 262 aa overlap
(232-462:316-567)

FIGURE 58C

```

      210      220      230      240      250
gi-50913506. KTTSLGSVLSYSLTFELPSYTKAEVNKTVYVSDNMSEGLTFNFNLSLTVEWKGMAN----
      |::: :|::: :: |: :|::: :
gi-19745307. EKEVAVDDAELKQINSEGOQEISVTWTNQLVTDE--KGMAYIYSVKEVDKNGELLEPKDY
      290      300      310      320      330      340

      260      270      280      290      300      310
gi-50913506. ITEDGSVMVENTKIGIAKEVNNGFNLSFIY-----DSLESISPNI SYKAVVNNKAIVGEE
      |::: :| || : |::: ::: : | |:::| |::: :| ||:
gi-19745307. IKKEDGLTVTNTYV---KPTSGHYDIEVTFGNHIDITEDTTPDI-VSGENQMKQIEGED
      350      360      370      380      390

      320      330      340      350
gi-50913506. GNP-----NKAEFFYSNNP-----TKGNTYDNL-DKKP-DKNGGITSKEDSKIVVYTYQ
      ::| | ||::: | |::| |::: | |::| |:::| |:::| |:::|
gi-19745307. SKPIDEV TENNLIEFGKNTMPGEEDGTNSNKYEEVEDSRPVDTL SGLSSEQQSGDMTIE
      400      410      420      430      440      450

      360      370      380      390      400      410
gi-50913506. -----IAFRKVD SVSKTFLIGAIFGVYDTSNKLIDIVTTNKNGYAISTQVSSGKYKIKE
      | | | |::: | ||: : |::| |::: | |:::| |:::| |:::|
gi-19745307. EDSATHIKFSKRD-IDGKELAGATMELRDSSGKTIS--TWISDGQVKDFYLMFGKYTFVE
      460      470      480      490      500      510

      420      430      440      450      460      470
gi-50913506. LKAPKGYSLNTE-TYEITAN-WVTATVKTSANSKSTTYTSDKNKATDNSEQVGWLKNGIF
      || ||: | |::: : ||::| |::: : ::: | | | | |
gi-19745307. TAAPDGYEIAITFTTVNEQGQVTVNGKATKGDHIVMV-DAYKPTKSGSQVIDIEEKLP
      520      530      540      550      560      570

      480      490      500      510      520      530
gi-50913506. YSIDSRPTGNDVK EAYIESTKALTDGTTFSKSNESGTVLLETDPNTKL GELPSTGSIG
gi-19745307. DEQGHSGSTTEIEDSKSSDVIIGGQGQIVETTEDTQTGMHGD SGCKTEVEDTKLVQSFHF
      580      590      600      610      620      630

```

gi-50913506.pep

/home/morana/gas/pili/align/gi-21909640.pep

gi|21909640|ref|NP_663908.1| protein F2 like fibronectin-binding protein [Streptococcus pyogenes MGAS315]

SCORES Init1: 56 Initn: 81 Opt: 134 z-score: 75.0 E(): 1.2
 >>/home/morana/gas/pili/align/gi-21909640.pep (656 aa)
 initn: 81 init1: 56 opt: 134 z-score: 75.0 expect(): 1.2
 Smith-Waterman score: 156; 23.9% identity in 347 aa overlap
 (232-547:202-524)

```

      210      220      230      240      250
gi-50913506. KTTSLGSVLSYSLTFELPSYTKAEVNKTVYVSDNMSEGLTFNFNLSLTVEWKGMAN----
      |::: :|::: :: |: :|::: :
gi-21909640. EKEVAVDDAELKQINSEGOQEISVTWTNQLVTDE--KGMAYIYSVKEVDKNGELLEPKDY
      180      190      200      210      220

      260      270      280      290      300      310
gi-50913506. ITEDGSVMVENTKIGIAKEVNNGFNLSFIY-----DSLESISPNI SYKAVVNNKAIVGEE
      |::: :| || : |::: ::: : | |:::| |::: :| ||:
gi-21909640. IKKEDGLTVTNTYV---KPTSGHYDIEVTFGNHIDITEDTTPDI-VSGENQMKQIEGED
      230      240      250      260      270      280

```

FIGURE 58D

FIGURE 58E

PCT/US2005/027239

```

gi-50913506. -----IAFRKVDSVSKTPLIGAIFGVYDTSNKLIDIVTTNKNGYAISTQVSSGKYKIKE
| | | | :: | | : : | : | | : | : : : : | | : |
gi-28810263. EDSATHIKFSKRD-IDGKELAGATMELRDSSGKTIS--TWISDGOVKDFYLMPGKYTFVE
460      470      480      490      500      510

      420      430      440      450      460      470
gi-50913506. LKAPKGYSINTE-TYEITAN-WVTATVKTSAKSKSTTYTSDKNKATDNSEQVGWLKNGIF
| | | : | | : : : | | : | : : : : : | | | : | |
gi-28810263. TAAPDGYEVATAITFTVNEQGQVTVNGKATKGDHIVMV-DAYKPTKSGSQV-----
520      530      540      550      560

      480      490      500      510      520      530
gi-50913506. YSIDSRTGNDVKEAYIESTKALTDGTTFSKSNEGSGTVLLETDPNTKLGLPSTGSIG
| | : | : : : | | : | : : : | : | : | : | : |
gi-28810263. --IDIEKLDP-EQGHSGSTTEIEDSKSSDVIIGGQGEVVDTE--DTQSGMTGHSGST-
570      580      590      600      610      620

      540      550
gi-50913506. TYLFKAIGSAAMIGAIGIYIVKRRKA
| : : : | : : : |
gi-28810263. TKIEDSKSSDVIVGGQGIIVETTEDTQTMHGDSGRKTEVEDTKLVQSFHFDNKEPESNS
630      640      650      660      670      680

```

gi-50913506.pep
/home/morama/gas/pili/align/orf80.pep

TRANSLATE of: orf80.seq check: 9824 from: 1 to: 1056
generated symbols 1 to: 352.
GETSEQ from morama, September 13, 2004 17:11.

SCORES Init1: 45 Initn: 69 Opt: 113 z-score: 70.8 E(): 2.1
>>/home/morama/gas/pili/align/orf80.pep (352 aa)
initn: 69 init1: 45 opt: 113 Z-score: 70.8 expect(): 2.1
Smith-Waterman score: 123; 22.8% identity in 311 aa overlap
(284-556:49-352)

```

      260      270      280      290      300      310
gi-50913506. KMANITEDGSVMVENTKIGIAKEVNNGFNLSFIYDSLESISPNISYKAVVN-NKAIVGEE
| | | | : | : | : | : : : : : | |
orf80.pep. ATALGTASLNQNVKAETAGVVTGKSLQVTKMTYDDEEVLMPEAFTFTIEPDMTASGKE
20      30      40      50      60      70

      320      330      340      350      360      370
gi-50913506. GNPN-KAEFFYSNNPTKGNITYDNLDRKPKDNGCITSKEDSKIVTYQIAFRKVDSVSKTP
| : : | : : : | | | : | : : : | : : : | : :
orf80.pep. GSLDTKNGIVEGLDKQVTVKYNKTDKPSQTK-LAQDFSKVKFPAIGVYRYMVSEKNDK
80      90      100      110      120      130

      380      390      400      410      420
gi-50913506. LIGAIFGVYDTSNKLIDIVTTNKNG-----YAISTQ-VSSGKYKIKELKAPKGYSIN
| : | : : : | : : | : | : | : : : | : : : | :
orf80.pep. KDGI---TYDDKKWTVDVYVGKANNNEEGFEVLYIVSKEGTSSTKKPIEFTNSIKTTSLK
140      150      160      170      180      190

      430      440      450      460
gi-50913506. TETYEITANW-----VTATVKTSAKSKSTTYTSDKNKATDNSEQVG-----WLKNGI
| : | : | : | : | : | : : : : : : : : : | :
orf80.pep. IEK-QITGNAGDRKKSFNFTLTLPSEYYKTGSVVKIEQDGSKKDVTIGTPYKFTLGHGK
200      210      220      230      240      250

470      480      490      500      510

```

FIGURE 58F

PCT/US05/27239

gi-50913506. FYSIDSRPTGNDV-----KEYAI-----ESTKALTDGTTFSKSNESGTVLLETDI
 :: | | : |:: | | : | | : | | : | | :
 orf80.pep SVMLSKLPIGINYLLSEDEANKDGYTTTATLKEQGKEKSSDFTLSTQNKTKDESADIVV
 260 270 280 290 300 310

520 530 540 550
 gi-50913506. PNTKLGELPSTGSIGTYLFKAIGSAAMIGAIGIYIVKRRKA
 | : :: | | : | | : | : | | : | | : | | : | | :
 orf80.pep TNKRDTQVP-TGVVGTLPAPFAVL SIVAIGGV-IYITKRKKA
 320 330 340 350

gi-50913506.pep
 /home/morاما/gas/pili/align/gi-19224137.pep

gi|19224137|gb|AAL86408.1|AF447492_5 EftLSL.A [Streptococcus pyogenes]

SCORES Initl: 45 Initn: 69 Opt: 109 z-score: 69.8 E(): 2.4
 >>/home/morاما/gas/pili/align/gi-19224137.pep (342 aa)
 initn: 69 initl: 45 opt: 109 Z-score: 69.8 expect(): 2.4
 Smith-Waterman score: 169; 26.0% identity in 334 aa overlap
 (257-556:25-342)

230 240 250 260 270 280
 gi-50913506. NKTVYVSDNMSEGLTFNFNLSLVEWKGKMANITEDGSVMVENTKIGIAKEVNNGFNLSFI
 | : : : | : : : | | : | |
 gi-19224137. MKKNKLLLATAILATALGTASLNQNVKAETAGVSSGQLTIKKSTTN-FN----
 10 20 30 40

290 300 310 320 330 340
 gi-50913506. YDSLESISPNI SYKAVVN-NKAIVGEEGN-PNKAEFFYSNNPTKGNTYDNLDDKPKDKNNG
 | : | : | : | | : | : | : | : | : | : | : | : | :
 gi-19224137. DDTL--LMPKTDYTFSVNPD SAATGTESNLPIKPGIAVNNQDIK-VSYNNTDKTSGKEKQ
 50 60 70 80 90 100

350 360 370 380 390 400
 gi-50913506. ITSKEDESKIVYTYQIAFRKVD SVSKTPLIGAIFGV-YDTSNKLIDIVTNNKNGYAISTQV
 : : : | : : : | : : | : | | : : | : : : : : : :
 gi-19224137. VVV-DFMKVTFPSVGIYRYVVTENK---GTAEGVTYDDTKWLVDVYVGNNEKGGLPEPKY
 110 120 130 140 150 160

410 420 430 440 450
 gi-50913506. SSGKYKIKELKAPKGY--SLNTETYBITANWVTATVKT SANSKSTTYTSDKNKATDNS--
 : | : | : | : | : | : | : | : | : | : | : | : | :
 gi-19224137. IVSKKGD SATKEPTQFNNSFTTSLKIEKE-VTCNTGDKKATFTTLTLPNEYEASSV
 170 180 190 200 210 220

460 470 480 490
 gi-50913506. ---EQVGWLKN---GIFYSI---DSR-----PTGND--VKEAYIE-----STKALTDG
 | : | : | : | : | : | : | : | : | : | : | : | : | :
 gi-19224137. VKIEENGQTRDVKIG EAYKFTLNDSQS SVLSKLPVGIN YKVEEAEANQGGYTTTATLKDG
 230 240 250 260 270 280

500 510 520 530 540 550
 gi-50913506. TTFSKSNEG---SGTVLLETDI PNTKLGELPSTGSIGTYLFKAIGSAAMIGAIGIYIVK
 : | | : : | : : : : | : | : | : | : | : | : | : | :
 gi-19224137. EKLSTYNLGQEHKTDKTADEIVVTNNRDTQVP-TGVVGTLPAPFAVL SIVAIGGV-IYITK
 290 300 310 320 330

gi-50913506. RRKA
 | : |

FIGURE 58G

PCT/US05/27239

194/487

gi-19224137. RKKA
340

gi-50913506.pep

/home/morama/gas/pili/align/gi-19224141.pep

gi|19224141|gb|AAL86412.1|AF447492_9 protein F2 [Streptococcus pyogenes]

SCORES Init1: 73 Initn: 73 Opt: 118 z-score: 68.9 E(): 2.7
>>/home/morama/gas/pili/align/gi-19224141.pep (1161 aa)
initn: 73 init1: 73 opt: 118 z-score: 68.9 expect(): 2.7
Smith-Waterman score: 174; 23.9% identity in 406 aa overlap
(115-483:277-645)

```
          90      100      110      120      130      140
gi-50913506. VSLTETGPTSQEITTIANGINTGKIKPFSTENVISISNGTATYNARGASVY--IALLTGAT
          ||:|||| |::: |::: |::: |
gi-19224141. IYTFDYYIAGLDKVLQSAELSLFLENKEVLENTSISNFKSTIGGQEITYKGTNVNVLGYNE
          250      260      270      280      290      300

          150      160      170      180      190
gi-50913506. DGRTYNPILLAASYNAGEGNLVTKNIDSKSNLYGQTSVAKSSLPSITKKVGTG-----T
          : : | | : | | : : | : : | : : : : | | : : |
gi-19224141. STKESNYITNGLSNVG-GSIESYNTETGEFVWYVYVNPRTNIPYATMNLWGFGRARSNT
          310      320      330      340      350      360

          200      210      220      230      240      250
gi-50913506. ID---DVNKKTTSLGSVLSYSITF--ELPSYTKAEVNKTFVVDNMSEGLTFNFNLSLTVE
          | | : | : : | : : | : | | | : | : : | : | : : |
gi-19224141. SDLENDANTSSAELGEIQVYEVPEGEKLPSSYGVDTVTKLTLRTD-ITAGLNGGFQ-----
          370      380      390      400      410

          260      270      280      290      300      310
gi-50913506. WKGKMANITEDGSVMVENTKIGIAKEVNNGFNLSFIYDSLESISPNI SYKAVVNNKAIVG
          | : : | : : | | | : | | : : : | | : | : : :
gi-19224141. -----MTKRQRIDFG--NNIQNKAFIIKV--TGKTDQSGKPLVVQSNLAS
          420      430      440      450      460

          320      330      340      350      360
gi-50913506. EEGNPNKAEFFYSNNPTKGNTRY--DNLDDKPKDKNGITSKEDSKIVYTY-----QIAF
          : | : | | : | : | : : : : | : | : | : | : | : |
gi-19224141. FRGASEYAAF-----TPVGGNVYFQNEIALSPSKGSGSGKSEETKPSITVANLKRVAQLRF
          470      480      490      500      510

          370      380      390      400      410      420
gi-50913506. RKVDSVSKTPLIGAFGVYDTSNKLIDI-VTINKNGYATSTQVSSGKYKIKELKAPKGYG
          : | : | : : | | : : : : : : : : : : | : | : | : |
gi-19224141. KKM-STDNVLPLEAFELRSSNGNSQKLEASSNTQGEVHFKDLTSGTYDLYETKAPKGYQ
          520      530      540      550      560      570

          430      440      450      460
gi-50913506. -----LNTETYEIT-----ANWVT--ATVKTSANSKSTTYTSDKNKATDNSEQVGWLKN
          | | | : | : : | : : : | | | : | : : | : : : | : :
gi-19224141. QVTEKLATVTVDTPKPAEMVTWGSPPHSSVKVEAN-KEVTIVNHKETLTFSGKKI-WEND
          580      590      600      610      620      630

          470      480      490      500      510      520
gi-50913506. GIFYSIDSRPTGNDVKEAYIESTKALTDGTTFSKSNESGTVLLETDPNTKIGELPSTG
          | : | : : | : :
gi-19224141. ----RPDQRFPAKIQVQLLQNGQKMPNQIQEVTKDNDWSYHFKDLPKYDAKNQYKYSVEE
          640      650      660      670      680
```

FIGURE 58H

gi-50913506.pep
/home/morama/gas/pili/align/gi-21909636.pep

gi|21909636|ref|NP_663904.1| conserved hypothetical protein [Streptococcus pyogenes MGAS315]

SCORES Init1: 45 Initn: 98 Opt: 96 z-score: 66.1 E(): 3.8
>>/home/morama/gas/pili/align/gi-21909636.pep (344 aa)
initn: 98 init1: 45 opt: 96 Z-score: 66.1 expect(): 3.8
Smith-Waterman score: 181; 25.3% identity in 312 aa overlap
(298-556:44-344)

```

      270      280      290      300      310      320
gi-50913506. NTKIGIAKEVNNGFNLSFIYDSLESISPNI SYKAVVNNKAIVGEEGNPNKAEFFYSNNPT
      :: : ::|::: : |::| ::: :
gi-21909636. ATALGTASLNQNVKAETAGVSENAKLIVKKTFSYTDNEVLMPKADYTFKVE---ADSTA
      20      30      40      50      60      70

      330      340      350      360
gi-50913506. KGNTYDNLDDKPKDGNGIT-----SKEDSKIVYTYQIAFRKV-----DSVSK
      :|:| |:|: || |:| :| |||: |::| || :||:
gi-21909636. SGKTKDGLEIKPGIVNGLTEQII SYTNTDKPDSKVKST-EFDFSKVVFPGIGVRYTVSE
      80      90      100      110      120

      370      380      390      400      410      420
gi-50913506. TPLIGAIFGV-YDTSNKLIDIVTTNKNGYAISTQVSSGKYKIKELKAPKGY--SINTETY
      | : | : ||:: :|: : ||:| :: : |: : ::| |:: |:| :
gi-21909636. KQ--GDVEGITYDTPKKWTVDVYVGNKEGGGFEPKFIVSKEQGTDVKKPVNFNNSFATTSL
      130      140      150      160      170      180

      430      440      450      460
gi-50913506. EITANWVTATVKT SAN-----SKSTTYTSDK-----NKATDNSE-QVGW-----LKNGI
      :: | | : : : ::||:: :|: :|::: :|::| |||
gi-21909636. KVKKNVSGNTGELQKEFDFTLT LNESTNFKKDQIVSLQKGNEKFEVKIGTPYKFKLKNGE
      190      200      210      220      230      240

      470      480      490      500      510
gi-50913506. FYSIDSRPTGNDVKEAYIESTKALTDGTTFSKS-NEGSGTVLL-----ETD-----
      ::|: |:| | :|::| || :|:|::| : :||
gi-21909636. SIQLDKLPVGITYKVNEMEANK---DGYKTTASLKEGPGQSKMYQLDMEQKTDESAD EIV
      250      260      270      280      290      300

      520      530      540      550
gi-50913506. IPNTIKLGELPSTGSIGTYLEKAIGSAAMIGAIGIYIVKRRKA
      : | : ::| || :|| :|: | : ||:: |||:|:|
gi-21909636. VTNKRDTQVP-TGVVGT LAPFAVL SIVAIGGV-IYITKRRKA
      310      320      330      340

```

gi-50913506.pep
/home/morama/gas/pili/align/gi-28810259.pep

gi|28810259|dbj|BAC63197.1| hypothetical protein [Streptococcus pyogenes SSI-1]

SCORES Init1: 45 Initn: 98 Opt: 96 z-score: 66.0 E(): 3.8
>>/home/morama/gas/pili/align/gi-28810259.pep (349 aa)
initn: 98 init1: 45 opt: 96 Z-score: 66.0 expect(): 3.8
Smith-Waterman score: 181; 25.3% identity in 312 aa overlap
(298-556:49-349)

FIGURE 58I

```

270      280      290      300      310      320
gi-50913506. NTKIGIAKEVNNGFNLSFIYDSLESISPNISYKAVVNNKAIVGEEGNPNKAEFFYSNNPT
                :: : ::|::: : |:: : :
gi-28810259. ATALGTASLNQNVKAETAGVSENAKLIVKKTFSYTDNEVLMPKADYTFKVE---ADSTA
20      30      40      50      60      70

330      340      350      360
gi-50913506. KGNTYDNLDDKKPKDGNGIT-----SKEDSKIVYTYQIAFRKV-----DSVSK
:|:| |:|: || |:| :| ||: | :| || :||:
gi-28810259. SGKTKDGLEIKPGIVNGLTEQIISYTNIDKPD SKVKST-EFDFSKVVFPGIGVYRYTVSE
80      90      100      110      120      130

370      380      390      400      410      420
gi-50913506. TPLIGAIFGV-YDTSNKLIDIVITNKNGYAISTQVSSGKYKIKELKAPKGY--SLNTETY
| : | : ||:: :|: : ||:| : : : | : : :| | : : | : | :
gi-28810259. KQ--GDVEGITYDTKKWTVDVYVGNKEGGGFEPKFIVSKEQGTVDVKKPVNFNNSFATTSL
140      150      160      170      180      190

430      440      450      460
gi-50913506. BITANVWTATVKTSAN-----SKSTTYTSDK-----NKATDNSE-QVGW-----LKNGI
: : | | : : : :|::: :|: :|::: | :| | |||
gi-28810259. KVKKNVSGNTGELQKEFDFTLTINESTNFKKDQIVSLQKGNKFEVKIGTPYKFKLKNGE
200      210      220      230      240      250

470      480      490      500      510
gi-50913506. FYSIDSRPTGNDVKEAYIESTKALTDGTTESKS-NEGSGTVLL-----ETD-----
:|:| |:| | :|:| || :| :|::| : :| :|
gi-28810259. SIQLDKLPVGITYKVNEMEANK---DGYKTTASLKEGDGQSKMYOLDMEQKTDSEADEIV
260      270      280      290      300

520      530      540      550
gi-50913506. IPNTKLGELPSTGSIGTYLFAIGSAAMIGAIGIYIVKRRKA
: | : :| ||:| | : | : ||:| ||:|:|
gi-28810259. VTNKRDTQVP-TGVVGTLPAPFAVL SIVAIGGV-IYITKRRKA
310      320      330      340

```

gi-50913506.pep
/home/morana/gas/pili/align/gas15.pep

GAS15 GAS15

SCORES Initl: 42 Initn: 68 Opt: 96 z-score: 63.8 E(): 5
>>/home/morana/gas/pili/align/gas15.pep (762 aa)
initn: 68 initl: 42 opt: 96 z-score: 63.8 expect(): 5
Smith-Waterman score: 96; 23.4% identity in 269 aa overlap
(283-535:222-470)

```

260      270      280      290      300
gi-50913506. GKMANITEDGSVMVENTKIGIAKEVNNGFNLSFIYDSLES-ISPNISYKA--VVNN--K
                ||:: ::|::: |:|::: | :| :
gas15.pep     VWYYS DNAPISNPDESFKRESESNLVSTS QLSLMRQALKQLIDPNLATKMPKQVPDDFQL
200      210      220      230      240      250

310      320      330      340      350      360
gi-50913506. AIVGEEGNPNKAEFFYSNNPTKGNTYDNLDDKKPKDGNGIT'SKEDSKIVYTYQIAFRK--V
:| | : :| : |:| :|: | | : : : | :| :
gas15.pep     SIFESDKGDYKNGYQNL LSGGLVPT---KPPTPGDPMPPNQPPQ---TTSVLIRKYAI
260      270      280      290      300

370      380      390      400      410      420

```

FIGURE 58J

FIGURE 58K

PCT/US05/27239
 gi-13621428. E-KTTKGGQAFVQT---EASIDQLYHFTLKDGESIKVTNLPVGVDYVVTTEDDYKSEKYT
 220 230 240 250 260

 500 510 520 530 540
 gi-50913506. T-----DGT-----FSKSNESGTVLLETDPNTKLGELPSTGSIGTYLFKAIGSAA
 | |::: | :| : : | | |::: :: : : |:| |
 gi-13621428. TNVEVSPQDGAVKNIAGNSTEQETSTDKDMTITFTNKKDFEVPTGVAMTVAPYIALGIVA
 270 280 290 300 310 320

 550
 gi-50913506. MIGAIGIYIVKRRKA
 : |: |:|:::|
 gi-13621428. VGGAL--YFVKKNA
 330 340

gi-50913506.pep
 /home/morاما/gas/pili/align/gi-19224135.pep

gi|19224135|gb|AAL86406.1|AF447492_3 Cpa [Streptococcus pyogenes]

SCORES Init1: 41 Initn: 41 Opt: 86 z-score: 61.0 E(): 7
 >>/home/morاما/gas/pili/align/gi-19224135.pep (756 aa)
 initn: 41 init1: 41 opt: 86 Z-score: 61.0 expect(): 7
 Smith-Waterman score: 101; 19.6% identity in 306 aa overlap
 (243-535:193-462)

 220 230 240 250 260 270
 gi-50913506. SLTFELPSYTKAEVNKTIVYVSDNMSEGLTFNFSNLTVEWKGKMANITEDGSVMVENTKIG
 ::: : | :: |:|: |:|:
 gi-19224135. PKNANGYMDKIEPLNAILVTQQAVWYYSDDSYGNIKTLWASEL----KDGKIDFEQVKL-
 170 180 190 200 210

 280 290 300 310 320 330
 gi-50913506. IAKEVNNGFNLSFIYDSLESISPNI SYKAVVNNKAIVGEEGNPN--KAEFFYSNNPTKGN
 :|: : ::| |:| | | :: | : ::: | |:|: : |: |:
 gi-19224135. -MREAYS----KLISDDLEETSKNKL PQGSKLNIFVPQDKSVQNLLSAEYVPESPPAPGQ
 220 230 240 250 260 270

 340 350 360 370 380
 gi-50913506. TYDNLDDKPKDNGGITSKEDSKIVYTY-QIAFRKVDSSVSKTPLIGAIFGVYDTSNKLIDI
 : | : |:| |:| | : : |: : | | |:| |:| |:
 gi-19224135. S---PEPP---VQTKTSVIIRKYAEGDYSKLLEGATLRLTGE--DILDFQEK--V
 280 290 300 310

 390 400 410 420 430 440
 gi-50913506. VTTNKNGYAISTQVSSGKYKIKELKAPKGYSLNTET-YETANWVTATVKTSA NSKSTT-
 :| | | ::|:| |:| |:| |:| |:| |:| |:| |:| |:
 gi-19224135. FQSNGTGEKI--ELSNGTYTLTETSSPDGYKIAEPIKFRVNNKKVFIVQKDG SQVENPNK
 320 330 340 350 360 370

 450 460 470 480 490
 gi-50913506. -----YTSKDNKATDNSEQVG---WLKNGIFYSIDSRPTGNDVKEAYIESTKALTDGTTF
 |: : : ::|: : : | | :: :::| : : | : |:
 gi-19224135. EVAEPYSVEAYSDMQDSNYINPETFTPYGKFYYAKNKKDKSSQVYCFN-----ADLHSP
 380 390 400 410 420 430

 500 510 520 530 540 550
 gi-50913506. SKSNEGSGTVLLETDPNTKLGELPSTGSIGTYLFKAIGSAAMIGAIGIYIVKRRKA
 :|:|:|:|: : |: | |: | : |:| |:
 gi-19224135. PESEDGGGTI--DPDISTMK--EVKYTHTAGSDFKYALRPRDTPEDFLKHIIKKVIEKG
 440 450 460 470 480

FIGURE 58L

199/487

gi-19224135. YNKKGDSYNGLTETQFRAATQLAIYYFTDSTDCLKTLKTYNNGKGYHGFESMDEKTLAVTK
490 500 510 520 530 540

gi-50913506.pep

/home/morana/gas/pili/align/gi-50913505.pep

gi|50913505|ref|YP_059477.1| Collagen adhesion protein [Streptococcus pyogenes M
GAS10394]

SCORES Init1: 70 Initn: 149 Opt: 86 z-score: 60.1 E(): 7.8
>>/home/morana/gas/pili/align/gi-50913505.pep (1036 aa)
initn: 149 init1: 70 opt: 86 z-score: 60.1 expect(): 7.8
Smith-Waterman score: 120; 21.5% identity in 469 aa overlap
(33-428:503-966)

gi-50913506.	NRRET	VREKILITAKK	MLACLA	ILAVVGL	CMTRVS	ALSKDD	TAQLKITNIEGG	PPTVTL	
gi-50913505.	YSQWVYHGT	FDYSSYQGES	FNRGQIH	YRYRGIV	SVSDGIR	RDDAVK	NSLLGVNG	-----L	
	480	490	500	510	520				
gi-50913506.	YKIGEG	VYNTNG	DSFINFK	----	YAE	GVSLTET	GTSGEIT	TIANGINTG	KIKPFSTEN
gi-50913505.	LQRFVN	INPENK	LSVIGF	QGSADY	HAGKWYP	DQSPRG	GFYQPNL	NNSRDAEL	LKGWSTNS
	530	540	550	560	570	580			
gi-50913506.	VSISNG	TATYNAR	GSVYIAL	ITGAT	-----	DGR	TYNPILLA	AS-----	YNGEGN
gi-50913505.	LLDENT	LTAHNN	GTNYHA	ALLKAKE	ILNEVK	DDGRR	KIMIFIS	DGVP	TFYFGED
	590	600	610	620	630	640			
gi-50913506.	NIDS--	KSNLYGQ	--TSVA	----	KSSL	PSITK	KVTGTID	DVNKKTT	SLGSVLS
gi-50913505.	NGSSND	RNNVTR	SQEGSK	LAIDF	KARYPN	LSIYSL	GVSKD	INSDT	ASSPVL
	650	660	670	680	690	700			
gi-50913506.	LPSY--	-----	TKEAVN	KIVY--	-----	VSDNM	SEGLTF	--NFNSL	TVEWKG
gi-50913505.	EHYYGI	TDPAEL	EKTENK	IVEDSK	LSQLGIS	DSLSQY	VDYD	KQPDV	LVTRK
	710	720	730	740	750	760			
gi-50913506.	-----	-----	DGSVM	VENTK	IGIAKE	VNNGFN	LSFIYDS	--LESIS	PNI----
gi-50913505.	ILYQK	DQVQE	AGKDI	DKVVT	TPKTT	SQPKG	VTLTF	FKSDY	KVDDEY
	770	780	790	800	810	820			
gi-50913506.	SYKAV	VNNKA	IVGEEG	NFNKA	EFFYS	NNPTK	GNTRYD	NLDKK	PKDK--
gi-50913505.	AYEKY	KDNEG	RYSEM	GSDT	TDYGT	NQTS	SGKGL	PSNSD	ASVNYM
	830	840	850	860	870	880			
gi-50913506.	IVYTY	QIAFR	KVDS	--VSKT	PLIGA	IFGVYD	TSNKLI	--DIVT	TNKN

FIGURE 58M

FIGURE 58N

```

! Distributed over 1 thread.
!   Start time: Wed Sep 15 18:45:54 2004
! Completion time: Wed Sep 15 18:46:02 2004
! CPU time used:
!   Database scan: 0:00:00.1
! Post-scan processing: 0:00:01.9
!   Total CPU time: 0:00:02.0
! Output File: gi-50913506.fasta

```

FIGURE 580

!!SEQUENCE LIST 1.0

(Peptide) FASTA of: gi-13621430.pep from: 1 to: 215 September 15, 2004 18:45

gi|13621430|gb|AAK33240.1| hypothetical protein [Streptococcus pyogenes]

TO: *.pep Sequences: 56 Symbols: 22,803 Word Size: 2

Scoring matrix: GenRunData:blosum50.cmp

Variable pamfactor used

Gap creation penalty: 12 Gap extension penalty: 2

Histogram Key:

Each histogram symbol represents 1 search set sequences:

z-scores computed from opt scores

z-score	obs	exp
(=)	(*)	
< 20	0	0:
22	0	0:
24	0	0:
26	0	0:
28	0	0:
30	0	0:
32	0	0:
34	0	1:*
36	0	2: *
38	0	3: *
40	0	4: *
42	0	5: *
44	1	5:= *
46	7	5:====*==
48	21	5:====*=====
50	7	5:====*==
52	2	4:== *
54	4	3:==*==
56	3	3:==*
58	0	2: *
60	4	2:==*==
62	1	2:==*
64	0	1:*
66	0	1:*
68	0	1:*
70	0	1:*
72	0	0:
74	0	0:
76	0	0:
78	0	0:
80	3	0:===
82	2	0:==
84	0	0:
86	0	0:
88	0	0:
90	0	0:
92	0	0:
94	0	0:
96	0	0:
98	0	0:
100	0	0:

FIGURE 59

102 PCT/US2005/027239
 104 0 0:
 106 0 0:
 108 0 0:
 110 0 0:
 112 0 0:
 114 0 0:
 116 0 0:
 118 0 0:
 >120 1 0:=-

Joining threshold: 36, opt. threshold: 24, opt. width: 16, reg.-scaled

The best scores are:

	init1	initn	opt	z-sc	E(55)...
/home/morama/gas/pili/align/gi-13621430.pep	Begin: 1	End: 215			
! gi 13621430 gb AAK33240.1 hypothet... 1338	1338	1338	233.9	1.8e-09	
/home/morama/gas/pili/align/gi-19745305.pep	Begin: 1	End: 193			
! gi 19745305 ref NP_606441.1 hypoth... 163	243	273	82.2	0.5	
/home/morama/gas/pili/align/gi-28810261.pep	Begin: 2	End: 187			
! gi 28810261 dbj BAC63199.1 hypothe... 164	239	268	81.5	0.55	
/home/morama/gas/pili/align/gi-19224139.pep	Begin: 2	End: 187			
! gi 19224139 gb AAL86410.1 AF447492... 164	236	265	81.0	0.57	
/home/morama/gas/pili/align/orf82.pep	Begin: 30	End: 222			
! TRANSLATE of: orf82.seq check: 4296... 163	235	264	81.0	0.58	
/home/morama/gas/pili/align/gi-21909638.pep	Begin: 2	End: 181			
! gi 21909638 ref NP_663906.1 hypoth... 164	239	261	80.5	0.62	
/home/morama/gas/pili/align/gi-19745303.pep	Begin: 84	End: 183			
! gi 19745303 ref NP_606439.1 hypoth... 121	121	126	61.4	6.7	
/home/morama/gas/pili/align/gi-13621428.pep	Begin: 6	End: 174			
! gi 13621428 gb AAK33238.1 hypothet... 58	86	122	60.9	7.2	
/home/morama/gas/pili/align/gi-19224137.pep	Begin: 93	End: 201			
! gi 19224137 gb AAL86408.1 AF447492... 88	88	119	60.4	7.5	
/home/morama/gas/pili/align/gi-50913503.pep	Begin: 549	End: 625			
! gi 50913503 ref YP_059475.1 Fibron... 73	73	117	60.4	7.6	
/home/morama/gas/pili/align/gi-19224134.pep	Begin: 631	End: 697			
! gi 19224134 gb AAL86405.1 AF447492... 73	73	115	60.1	7.8	

\\End of List

gi-13621430.pep

/home/morama/gas/pili/align/gi-13621430.pep

gi|13621430|gb|AAK33240.1| hypothetical protein [Streptococcus pyogenes]

SCORES Init1: 1338 Initn: 1338 Opt: 1338 z-score: 233.9 E(): 1.8e-09
 >>/home/morama/gas/pili/align/gi-13621430.pep (215 aa)
 initn: 1338 init1: 1338 opt: 1338 Z-score: 233.9 expect(): 1.8e-09
 Smith-Waterman score: 1338; 100.0% identity in 215 aa overlap
 (1-215:1-215).

gi-13621430.	10	20	30	40	50	60
	MKKSILRILAIGYLLMSFCLLDSVEAENLTASINIEVINQVDVATNKQSSDIDETFMFVI					
gi-13621430.	MKKSILRILAIGYLLMSFCLLDSVEAENLTASINIEVINQVDVATNKQSSDIDETFMFVI					
	10	20	30	40	50	60
gi-13621430.	70	80	90	100	110	120
	EALDKESPLPNSVTTTSVKNGKTSFEQLTFSEVGQYHYKIHQLLGKNSQYHYDETVYEVV					
gi-13621430.	EALDKESPLPNSVTTTSVKNGKTSFEQLTFSEVGQYHYKIHQLLGKNSQYHYDETVYEVV					

FIGURE 59A

FIGURE 59B

(8-213-2-187) US 05/27239

```

      10      20      30      40      50      60
gi-13621430. MKKSILRILAIGYLLMSFCLLDSEVAENLTASINIEVINQVDVATNKQSSDIDETFMFVI
      :::: ::| : : ::| |:: |:: :| | | :: | :: | ::
gi-28810261.  MLFSVVMILTMLAFNQTVLAKDSTVQTSISVENVLERAGDSTP-----FSIAL
      10      20      30      40

      70      80      90      100     110     120
gi-13621430. EALDKESPLPNSVTTTSVKNGKTSFEQLTFSEVGQYHYKIHQLLGKNSQYHYDETVEYEVV
      |::| : : : : :: |::|::| |::| |::| |::| :::|::| : | |::|:
gi-28810261. ESIDAMKTIEE---ITIAGSGKASESPLTFTTVGQYTYRVYQKPSQNKDYQADTTVFVDVL
      50      60      70      80      90      100

      130     140     150     160     170     180
gi-13621430. IYVLYNEQSGALETNLVSNNKLGETEKSELIFKQEYSEKTPEPHQPDTEKEKPKQKRNGI
      :|| |::| :| :::| : |::| : || : : | |::| : |::| :
gi-28810261. VYVTYDED-GTLVAKVISRRAGDEEKSAITFKPKWLKPIPPROPNI-----PKTP----
      110     120     130     140     150

      190     200     210
gi-13621430. LPSTGEMVSYVSALGIVLVATITLYSIYKKLKTSK
      || :||: | : : |::|::| : :| : |||:
gi-28810261. LPLAGEVKSLGILSIVLLGLLVLLYV-KKLSKRL
      160     170     180

```

gi-13621430.pep

/home/morama/gas/pili/align/gi-19224139.pep

gi|19224139|gb|AAL86410.1|AF447492_7 unknown [Streptococcus pyogenes]

```

SCORES   Init1: 164   Initn: 236   Opt: 265   z-score: 81.0   E(): 0.57
>>/home/morama/gas/pili/align/gi-19224139.pep (189 aa)
initn: 236 init1: 164 opt: 265 Z-score: 81.0 expect(): 0.57
Smith-Waterman score: 303;   30.6% identity in 206 aa overlap
(8-213:2-187)

```

```

      10      20      30      40      50      60
gi-13621430. MKKSILRILAIGYLLMSFCLLDSEVAENLTASINIEVINQVDVATNKQSSDIDETFMFVI
      :::: ::| : : ::| |:: |:: :| | | :: | :: | ::
gi-19224139.  MLFSVVMILTMLAFNQTVLAKDSTVQTSISVENVLERAGDSTP-----FSIAL
      10      20      30      40

      70      80      90      100     110     120
gi-13621430. EALDKESPLPNSVTTTSVKNGKTSFEQLTFSEVGQYHYKIHQLLGKNSQYHYDETVEYEVV
      |::| : : : : :: |::|::| |::| |::| |::| :::|::| : | |::|:
gi-19224139. ESIDAMKTIEE---ITIAGSGKASESPLTFTTVGQYTYRVYQKPSQNKDYQADTTVFVDVL
      50      60      70      80      90      100

      130     140     150     160     170     180
gi-13621430. IYVLYNEQSGALETNLVSNNKLGETEKSELIFKQEYSEKTPEPHQPDTEKEKPKQKRNGI
      :|| |::| :| :::| : |::| : || : : | |::| : |::| :
gi-19224139. VYVTYDED-GTLVAKVISRRAGDEEKSAITFKPKRLVKPIPPROPNI-----PKTP----
      110     120     130     140     150

      190     200     210
gi-13621430. LPSTGEMVSYVSALGIVLVATITLYSIYKKLKTSK
      || :||: | : : |::|::| : :| : |||:
gi-19224139. LPLAGEVKSLGILSIVLLGLLVLLYV-KKLSKSL
      160     170     180

```

FIGURE 59C

PCT/US05/27239

gi-13621430.pep
/home/morama/gas/pili/align/orf82.pep

TRANSLATE of: orf82.seq check: 4296 from: 1 to: 672
generated symbols 1 to: 224.

GETSEQ from morama, September 13, 2004 17:09.

SCORES Init1: 163 Initn: 235 Opt: 264 z-score: 81.0 E(): 0.58
>>/home/morama/gas/pili/align/orf82.pep (224 aa)
initn: 235 init1: 163 opt: 264 Z-score: 81.0 expect(): 0.58
Smith-Waterman score: 304; 30.5% identity in 213 aa overlap
(1-213:30-222)

```

gi-13621430.      10      20      30
                  MKKSILRILAIGYLLMSFCLLDSVEAENLTA
orf82.pep      LLFQVRKIFLLTIVLSLSVLFKNNERRRLRKYWKMLFSVVMILTMLAFNQTVLAKDSTV
                  10      20      30      40      50      60

gi-13621430.      40      50      60      70      80      90
                  SINIEVINQVDVATNKQSSDIDETFMFVIEALDKESPLNSVTSVKGNKGTSTFEQLTFFS
orf82.pep      QTSISVENVLERAGDSTP-----FSVALESIDAMKTI-DEIT--IAGSGKASFSPLTFT
                  70      80      90      100     110

gi-13621430.      100     110     120     130     140     150
                  EVGQYHYKIHQLLGKNSQYHYDETVEYEVVIYVLYNEQSGALETNLVSNKLGETEKSELIF
orf82.pep      TVGQYTYRVYQKPSQNKDYQADTTVFVDLVVYTYDED-GTLVAKVISRRAGDEEKSAITF
                  120     130     140     150     160     170

gi-13621430.      160     170     180     190     200     210
                  KQEYSEKTPEPHQPDTEKEKPKKRNGILPSTGEMVSYVSALGIVLVATITLYSIYKKL
orf82.pep      KPKRLVKPIPPROPNI-----PKTP-----LPLAGEVKSLLGILSIVLLGLLVLLYV-KKL
                  180     190     200     210     220

gi-13621430.      KTSK
orf82.pep      KSRL

```

gi-13621430.pep
/home/morama/gas/pili/align/gi-21909638.pep

gi|21909638|ref|NP_663906.1| hypothetical protein [Streptococcus pyogenes MGAS31
5]

SCORES Init1: 164 Initn: 239 Opt: 261 z-score: 80.5 E(): 0.62
>>/home/morama/gas/pili/align/gi-21909638.pep (183 aa)
initn: 239 init1: 164 opt: 261 Z-score: 80.5 expect(): 0.62
Smith-Waterman score: 302; 31.5% identity in 200 aa overlap
(14-213:2-181)

```

gi-13621430.      10      20      30      40      50      60
                  MKKSILRILAIGYLLMSFCLLDSVEAENLTASINIEVINQVDVATNKQSSDIDETFMFVI
gi-21909638.      :| : : :| : : :| : : :| : : :| : : :| : : :
                  MILTMLAFNQTVLAKDSTVQTSISVENVLERAGDSTP-----FSIAL

```

FIGURE 59D

PCT/US05/27239

```

                70      80      90      100      110      120
gi-13621430.  EALDKESPLPNSVTTSVKGNKTSFEQLTFSEVGQYHYKIHQLLGKNSQYHYDETVEV
|::| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
gi-21909638.  ESIDAMKTIEE---ITIAGSGKASFSPLTFTTVGQYTYRVYQKPSQNKDYQADTTVFDVL
                50      60      70      80      90

                130      140      150      160      170      180
gi-13621430.  IYVLYNEQSGALETNLVSNNKLGETEKSELIFKQYSEKTPPEHPQDPTTEKEKPQKKRNGI
:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
gi-21909638.  VYVTYDED-GTLVAKVISRRAGDEEKSAITFKPKWLVKPIPPROPNI-----PKTP----
                100      110      120      130      140

                190      200      210
gi-13621430.  LPSTGEMVSIVSALGIVLVATITLYSIYKLLKTSK
|| :||: | :| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
gi-21909638.  LPLAGEVKSLGILSIVLLGLLVLLYV-KKLSRL
                150      160      170      180

```

gi-13621430.pep

/home/morana/gas/pili/align/gi-19745303.pep

gi|19745303|ref|NP_606439.1| hypothetical protein [Streptococcus pyogenes MGAS8232]

```

SCORES  Init1: 121  Initn: 121  Opt: 126  z-score: 61.4  E(): 6.7
>>/home/morana/gas/pili/align/gi-19745303.pep (344 aa)
  initn: 121 init1: 121 opt: 126 Z-score: 61.4 expect(): 6.7
Smith-Waterman score: 126; 27.0% identity in 100 aa overlap
(59-155:84-183)

```

```

                30      40      50      60      70      80
gi-13621430.  LTASINIEVINQVDVATNKQSSDIDETFMFVIEALDKESPLP--NSVTTSVKNGN-KTSF
|::| : : : : : : : : : : : : : : : : : : : : : : : :
gi-19745303.  LMPKADYTFKVEADDNAKGKTKDGLDIKPGVIDGLENTKTIHYGNSDKTTAKEKSVNFD
                60      70      80      90      100      110

                90      100      110      120      130      140
gi-13621430.  EQLTFSEVGQYHYKIHQLLGKNSQYHYDETVEVVIYVLYNEQSGALETNLVSNNKLGETE
:|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
gi-19745303.  ANVKFPGVGVIYRTVSEVNGNKAGIAYDSQQWTVDVYVNNREDGGFEAKYIVSTEGGQSD
                120      130      140      150      160      170

                150      160      170      180      190      200
gi-13621430.  KSELIFKQYSEKTPPEHPQDPTTEKEKPQKKRNGILPSTGEMVSIVSALGIVLVATITLY
|: :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
gi-19745303.  KKPVLFRNFFDITSLKVTKKVTGNTGEHQRSFSFTLLLPNECFEKGQVNNILOGETKK
                180      190      200      210      220      230

```

gi-13621430.pep

/home/morana/gas/pili/align/gi-13621428.pep

gi|13621428|gb|AAK33238.1| hypothetical protein [Streptococcus pyogenes]

```

SCORES  Init1: 58  Initn: 86  Opt: 122  z-score: 60.9  E(): 7.2
>>/home/morana/gas/pili/align/gi-13621428.pep (340 aa)
  initn: 86 init1: 58 opt: 122 Z-score: 60.9 expect(): 7.2
Smith-Waterman score: 135; 29.1% identity in 172 aa overlap

```

FIGURE 59E

(8-159) PCT/US2005/027239

```

      10      20      30      40      50
gi-13621430. MKKSILRILAIGYLLMSFCLLDSVEAENLTASINIEVINQVD-VATNKQSSDIDETFMF-
      :| | | | :|::|::| : : | : :| | : | : | :
gi-13621428. MKLRHLLLTGAALTSFAAT-TVHGETVVNGAKLTVTKNLDLVNSNALIPNTDFTFKIE
      10      20      30      40      50

      60      70      80      90      100
gi-13621430. ---VIEALDK-----ESPLPN-SVTTSVKNGKTSFEQLTFSEV-----GQYHYKI-H
      | | :| :|::|::| :| | |::|: : : ||| | |::| :
gi-13621428. PDTTVNEDGNKFKGVALNTPMTKVITYNSDKGGSNTKTAEFDFSEVTFEKPQGVYVYKVT
      60      70      80      90      100      110

      110      120      130      140      150      160
gi-13621430. QLLGKNSQYHYDETVYEVVYVLYNE-QSCALETNLVSNKLGETEKSELIFKQEYSEKTP
      : : | | | | :|::|::| : : | : | : | : | : | :
gi-13621428. EKIDKVPQVSYDTSYTVQVHVLWNEEQQKPVATYIVGYKEGS--KVPIQFKNSLDSTTL
      120      130      140      150      160      170

      170      180      190      200      210
gi-13621430. EPHQPDTEKEKPKQKRNGLPSTGEMVSYVSALGIVLVATITLYSIYKKLKTSK
gi-13621428. TVKKKVSQGGDRSKDFNGLTLKANQYKASEKVMIEKTTKGGQAPVQTEASIDQLYHF
      180      190      200      210      220      230

```

gi-13621430.pep

/home/morama/gas/pili/align/gi-19224137.pep

gi|19224137|gb|AAL86408.1|AF447492_5 EftLSL.A [Streptococcus pyogenes]

```

SCORES   Initl: 88   Initn: 88   Opt: 119   z-score: 60.4   E(): 7.5
>>/home/morama/gas/pili/align/gi-19224137.pep   (342 aa)
initn: 88 initl: 88 opt: 119 Z-score: 60.4 expect(): 7.5
Smith-Waterman score: 119;   29.7% identity in 111 aa overlap
(72-176:93-201)

```

```

      50      60      70      80      90
gi-13621430. DVATNKQSSDIDETFMFVIEALDKESPLPNVTTSVKNGK-----TSFEQLTFSEVGQY
      | : : | : : | : : | : : | : : | : : | : : |
gi-19224137. SVNPDSAATGTESNLPKPGIAVNNQDIKVSYSNTDKTSGKEKQVVVDPMKVTFFPSVGIY
      70      80      90      100      110      120

      100      110      120      130      140      150
gi-13621430. HYKIHOLLGKNSQYHYDETVYEVVYVLYNEQSCALETNLVSNKLGETEKSELIFKQEY
      : | : : | : : | : : | : : | : : | : : | : : |
gi-19224137. RYVVTENKGTAEQVYDDTKVLVDVYVGNNEKGG-LEPKYIVSKKGD SATKEPIQFNNSF
      130      140      150      160      170      180

      160      170      180      190      200      210
gi-13621430. SEKTPEPHQPDTEKEKPKQKRNGLPSTGEMVSYVSALGIVLVATITLYSIYKKLKTSK
      | | : : | : : | : : | : : | : : | : : |
gi-19224137. -ETTSLKIEKEVTGNTGDHKAFTFTLTLPNEYEASSVVKIEENGQTKDVKICEAYKF
      190      200      210      220      230      240

```

gi-13621430.pep

/home/morama/gas/pili/align/gi-50913503.pep

gi|50913503|ref|YP_059475.1| Fibronectin-binding protein [Streptococcus pyogenes MGAS10394]

FIGURE 59F

PCT/US05/27239

SCORES Init1: 73 Initn: 73 Opt: 117 z-score: 60.4 E(): 7.6
 >>/home/morana/gas/pili/align/gi-50913503.pep (627 aa)
 initn: 73 init1: 73 opt: 117 Z-score: 60.4 expect(): 7.6
 Smith-Waterman score: 118; 28.7% identity in 87 aa overlap
 (129-215:549-625)

```

      100      110      120      130      140      150
gi-13621430. KIHQLLGKNSQYHYDETVEVVIYVLYNEQSGALETNLVSNKLGETEKSELIFKQEYSEK
                        || || | | : : | : : : |
gi-50913503. IETEDTKEPEVLMGGQSESVEFTKDTQTGMSGFSETATVV---EDTRPKLVFHFDDNNEP
      520      530      540      550      560      570

      160      170      180      190      200      210
gi-13621430. TPEPHQPDTEKEKPKQKRNGILPSTGEMVSYVSALGIVLVATITLYSIYKKLKTSK
      | : : : || | : : || : || : : : : || : : : : : : | : : : |
gi-50913503. KVEEN---REKPTKNITPILPATGDIENVLAFLGILILSVLSIFSLLKNKQSNKKV
      580      590      600      610      620

```

gi-13621430.pep

/home/morana/gas/pili/align/gi-19224134.pep

gi|19224134|gb|AAL86405.1|AF447492_2 protein F [Streptococcus pyogenes]

SCORES Init1: 73 Initn: 73 Opt: 115 z-score: 60.1 E(): 7.8
 >>/home/morana/gas/pili/align/gi-19224134.pep (698 aa)
 initn: 73 init1: 73 opt: 115 Z-score: 60.1 expect(): 7.8
 Smith-Waterman score: 115; 27.4% identity in 73 aa overlap
 (143-215:631-697)

```

      120      130      140      150      160      170
gi-13621430. DETVEVVIYVLYNEQSGALETNLVSNKLGETEKSELIFKQEYSEKTPEPHQPDTEKEK
                        | : : | : : : | | : : ||
gi-19224134. VLMGGQSESVEFTKDTQTGMSGFSETVTIVEDTRPKLVFHFDDNNEPKVEEN---REK
      610      620      630      640      650

      180      190      200      210
gi-13621430. PKQKRNGILPSTGEMVSYVSALGIVLVATITLYSIYKKLKTSK
      | | : || : || : : : || : : : : : : | : : : |
gi-19224134. PTKNITPILPATGDIENVLAFLGILILSVLSIFSLLKNKQNNKV
      660      670      680      690

```

! Distributed over 1 thread.

! Start time: Wed Sep 15 18:45:36 2004

! Completion time: Wed Sep 15 18:45:38 2004

! CPU time used:

! Database scan: 0:00:00.1

! Post-scan processing: 0:00:00.3

! Total CPU time: 0:00:00.4

! Output File: gi-13621430.fasta

FIGURE 59G

TYPE 3 pilus motifs
protein F2 like fibronectin-binding protein
Length: 696-733
LPXTG
pilin motif consensus PK (X₇) K
E box consensus ETxAPxGY

SpyM3_0104/21909640	pilin motif	155 PKEKPIIYFK
	E box	398 YTFVETAAPDGY
SPs0106/28895018	pilin motif	269 PKEKPIIYFK
	E box	512 YTFVETAAPDGY
SpyM18_0132/19745307	pilin motif	269 PKEKPIIYFK
	E box	512 YTFVETAAPDGY
orf84	pilin motif	269 PKEKPIIYFK
	E box	512 YTFVETAAPDGY

TYPE 4 pilus motifs
protein F2 like fibronectin-binding protein
Length: 1161
LPXTG
pilin motif consensus PK (X₇₋₈) K
E box consensus YxLxETxAPxGY

The protein is longer than the proteins belonging to TYPE 3 and has 4 possible pilin motifs and 2 E boxes

19224141	pilin motifs	215 PKGISQDIPK
		571 PKGYQQVTEK
		156 PKMSVVSKYK
		674 PKYDAKNQEVK
	E boxes	563 YDLYETKAPKGY
		940 YTFVETAAPDGY

FIGURE 60

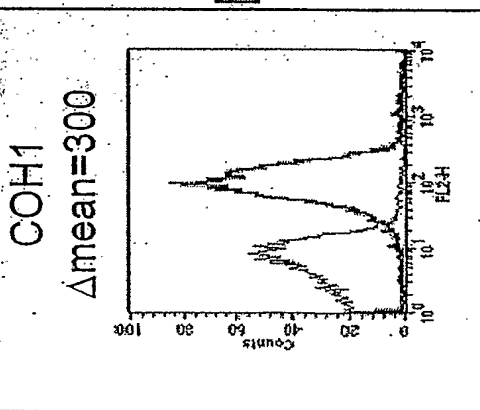
Formation of pili structures on GBS appears to be correlated to FACS values for surface expression of GBS80 protein

WO 2006/078318

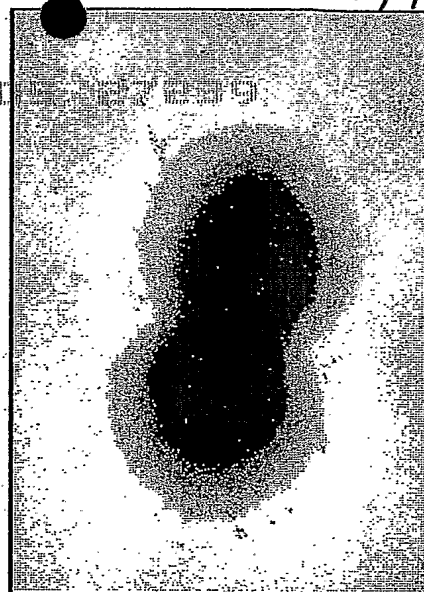
PCT/US2005/027239

211/487

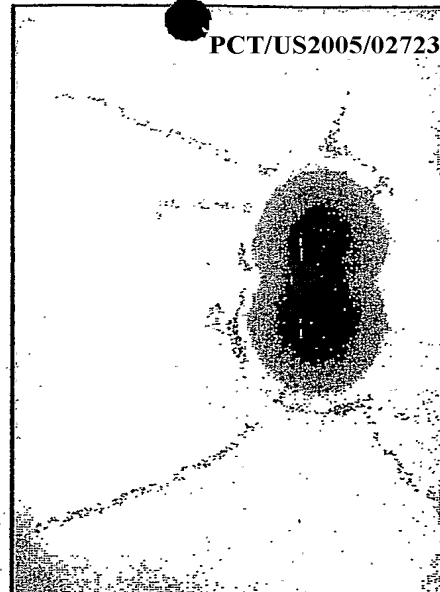
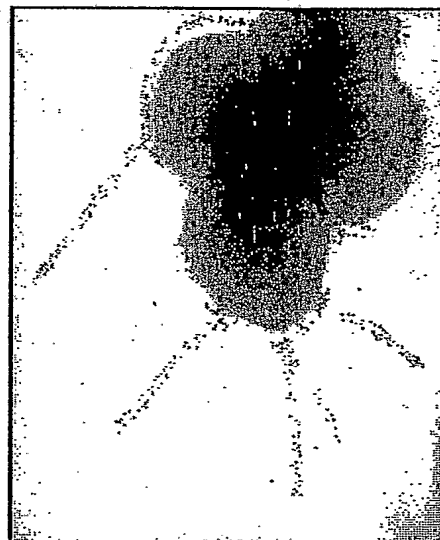
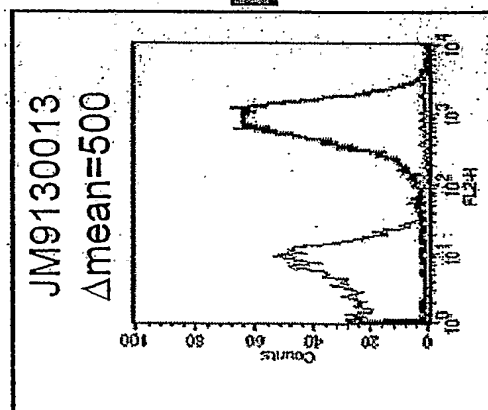
α -GBS80



Immunogold Electron Microscopy



α GBS80



PCT/US2005/027239

CHIRON VACCINES

FIGURE 61

Surface exposure is capsule-dependent for GBS 322 but not for GBS 80

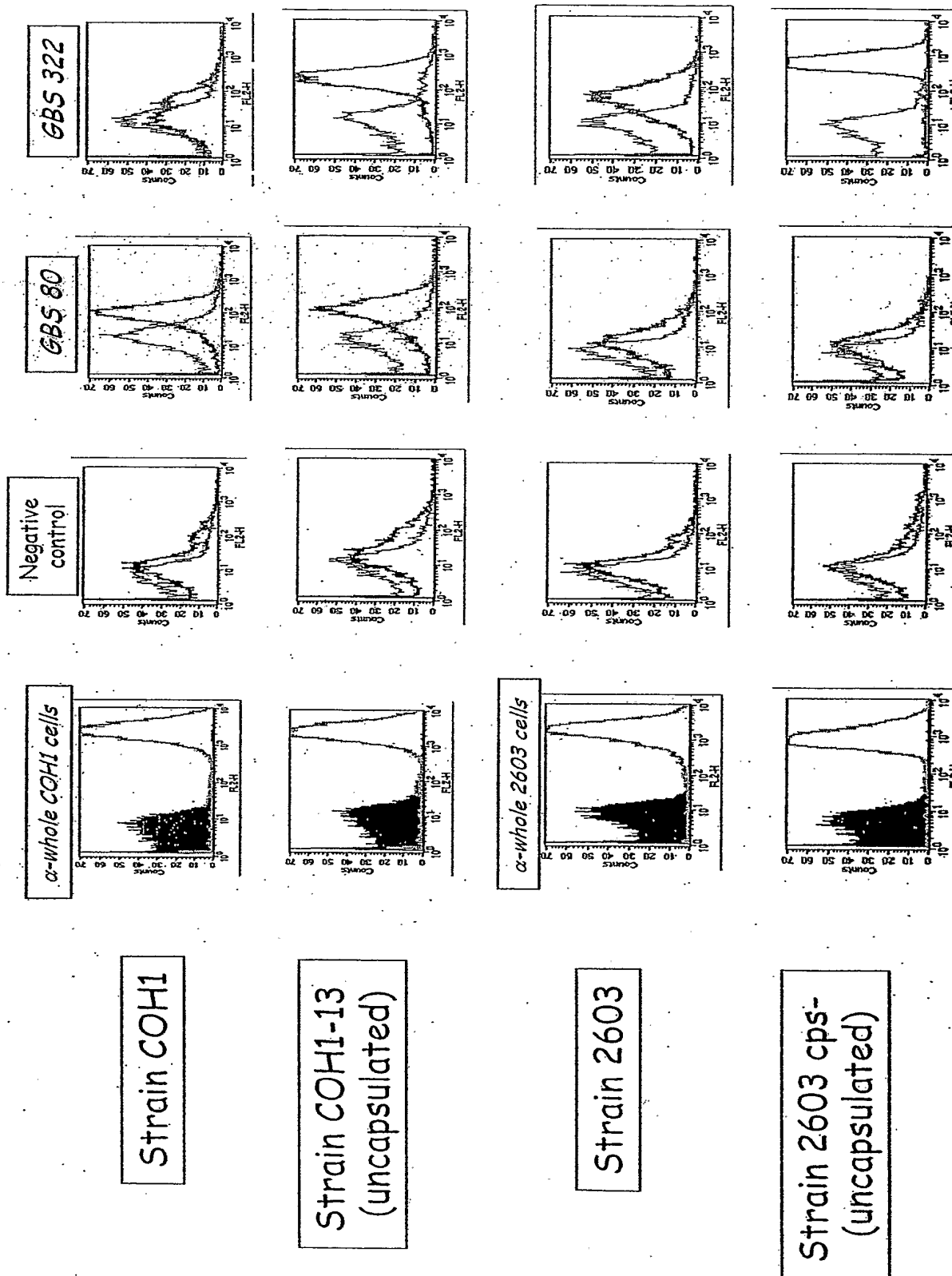
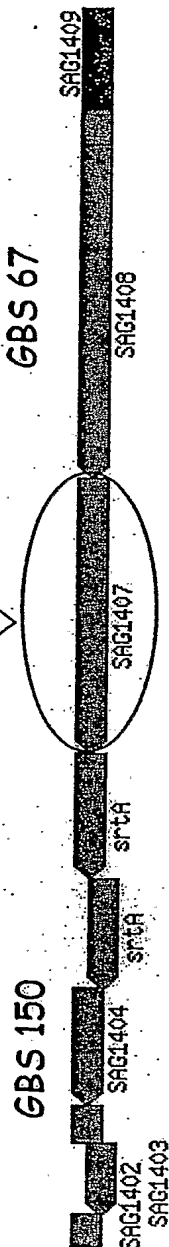


FIGURE 62

Adhesin island 2-

Operon gbs 67, 59, 150



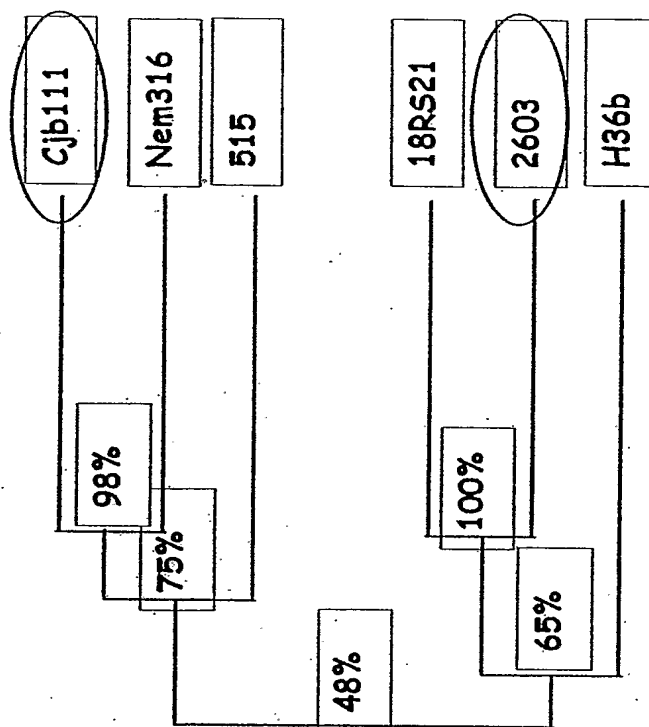
• 30 % identity with GBS 80

• By comparing GBS 59 amino acid sequence of 2603 with that of other sequenced genomes, the following homologies are obtained:

2603	
nem316	47%
h36b	62%
515	48%
cjb111	48%
18rs21	100%
coh1	not present (Spb1)
A909	not present (Spb1)

• CGH: 1/20 GBS strains analyzed (18RS21)

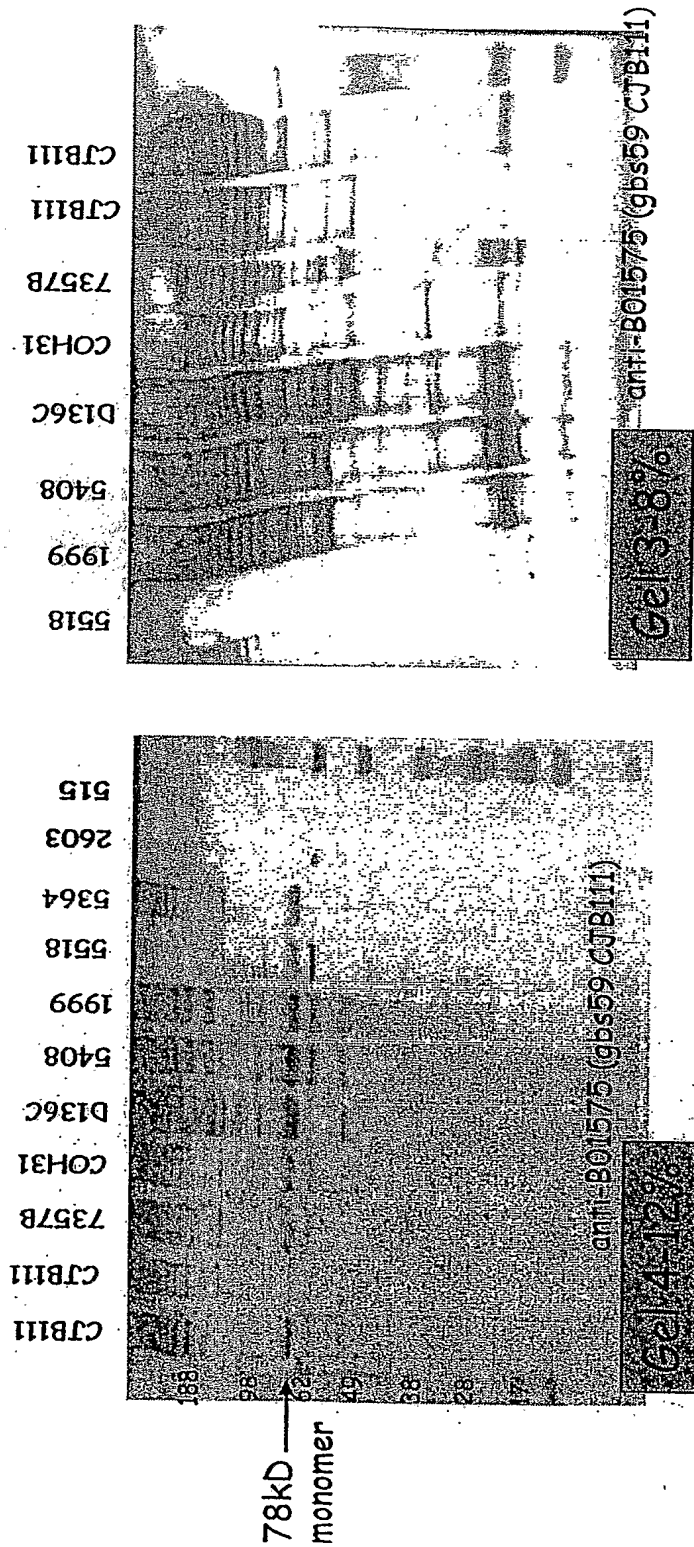
Two-by-two amino acid sequence comparison



There seem to be two clusters suggesting the presence of two major isoforms

FIGURE 63

Western blotting with whole extracts derived from GBS strains



GBS 59 is part of a high molecular weight polymer (pilus)
in GBS strains: CJB111, 7357b, coh31, d1363c, 5408, 1999,
5364, 5518, 515

FIGURE 64

Western blotting with purified proteins and whole extracts derived from GBS strains

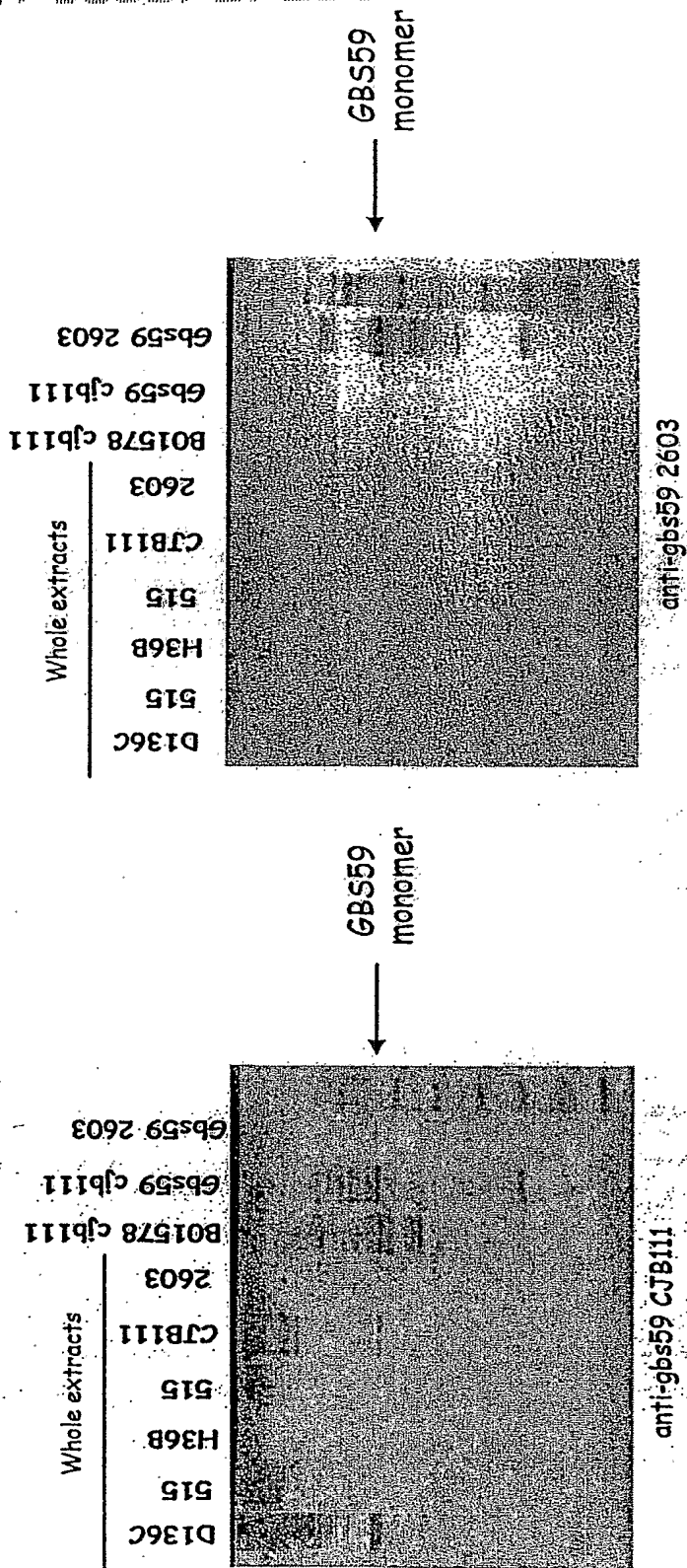


FIGURE 65

FACS analysis using mouse antiserum after immunization with BO1575 (gbs59) from CJB111 genome

WO 2006/078318

216/487

PCT/US2005/027239

PCT/US05/27239

GBS strains	Type	GBS 59
DK1	Ia	565
DK8		559
Davis		577
515		583
2986		443
5551		524
7357b-		596
5518		190
D136C		504
COH31	III	505
DK21	II	249
CJB111	V	493
5364		593
2110		590
1999		594
2210		636
5408		537
1169		227

GBS 59

GBS 67

PBS

CJB111

7357B

515

Where present GBS 59 is a highly exposed on the GBS surface

FIGURE 66

Opsonophagocytosis assays: B01575 (gbs59-cjb111)

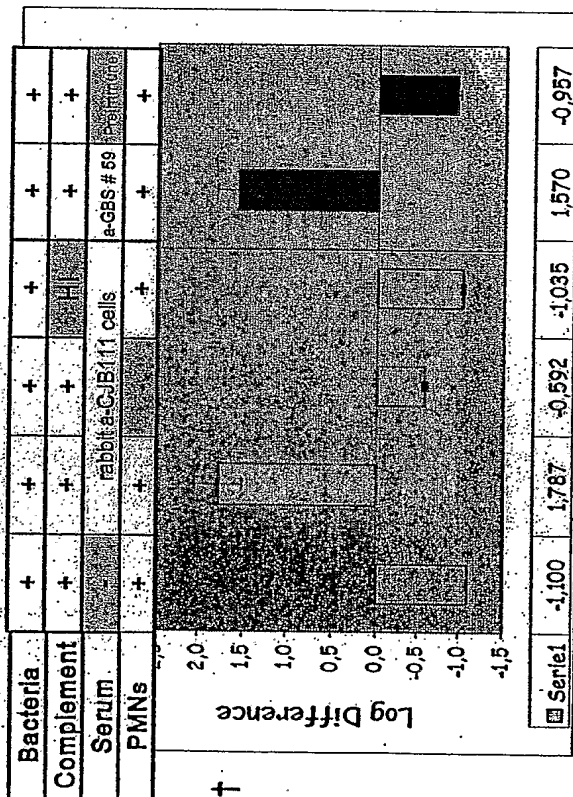
WO 2006/078318

PCT/US05/27239

217/487

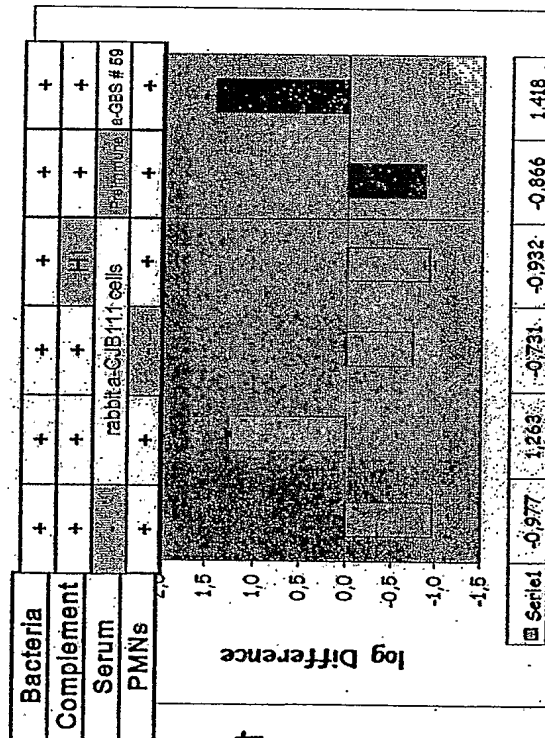
PCT/US2005/027239

I
experiment



- CJB111 GBS strain type Ia
- Baby rabbit complement
- Human PMNs
- Positive control: anti-type V cells (rabbit serum anti fixed type V cells)

II
experiment



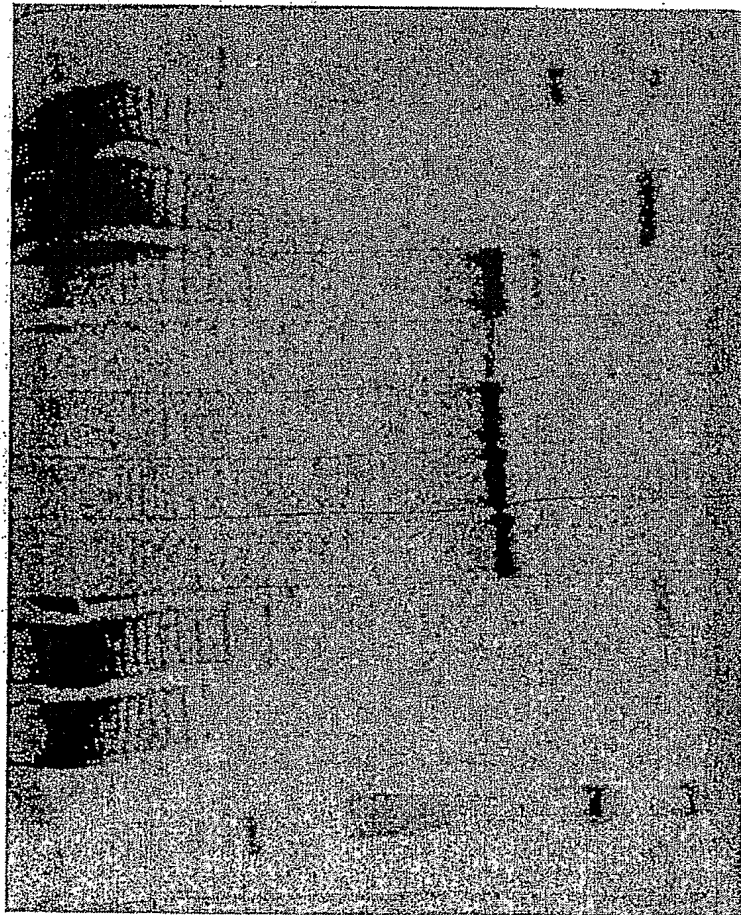
Antibodies against B01575 (cjb111) are opsonic for cjb111 GBS strain serotype V

FIGURE 67

Association GBS 80-104

WB α -80, α -104 JM9130013 Total Extract

Mab α -80 A4/77
 Mab α -80 19G4/78
 Mab α -104 15H3/49
 Mab α -104 12A7/67
 Mab α -104 H2/32
 Mab α -104 14B3/73
 α -104 POLIC.
 α -80 POLIC.
 Mab α -80 19G4/77

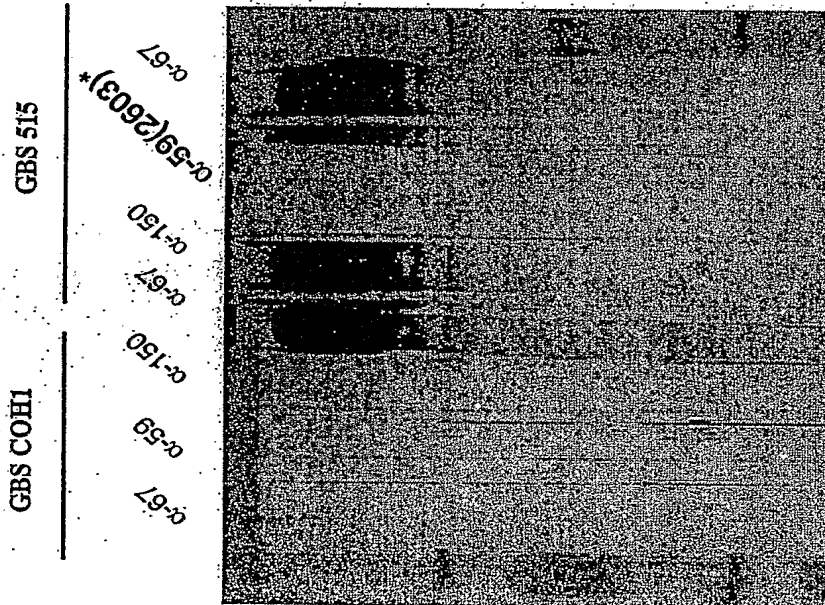


FACS	(Δ MEAN)
GBS 80	597
GBS 104	446

FIGURE 68

WB GBS 515 Total Extract

α -67; α -150; α -59 (2603)



Controls:
GBS COH1 total extracts

- anti-gbs59 mouse serum after immunization with SAG1407 (GBS 59) from 2603 genome

GBS 67 and GBS 150 are parts of a high molecular weight polymer (pilus) in 515 GBS strain

FIGURE 69

Western Blotting ko GBS67 from 515 genome (clone 1.45)

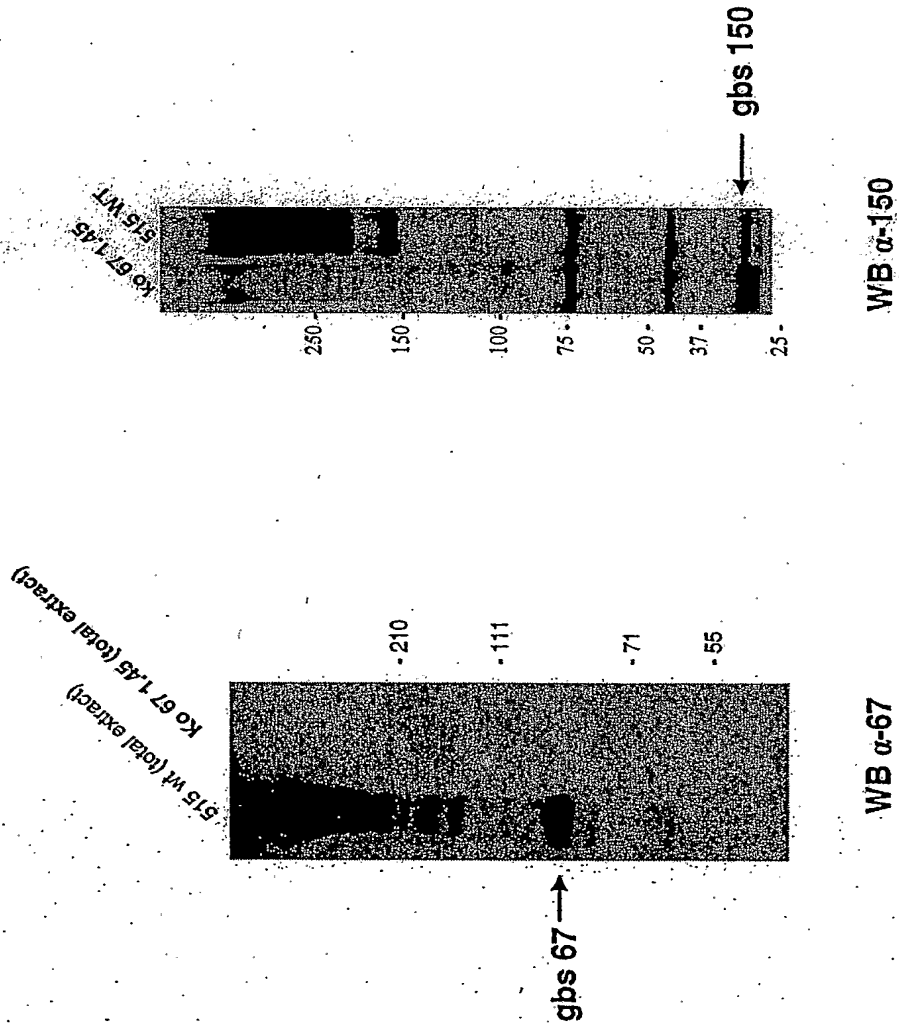


FIGURE 70

FACS GBS 515 Δ67

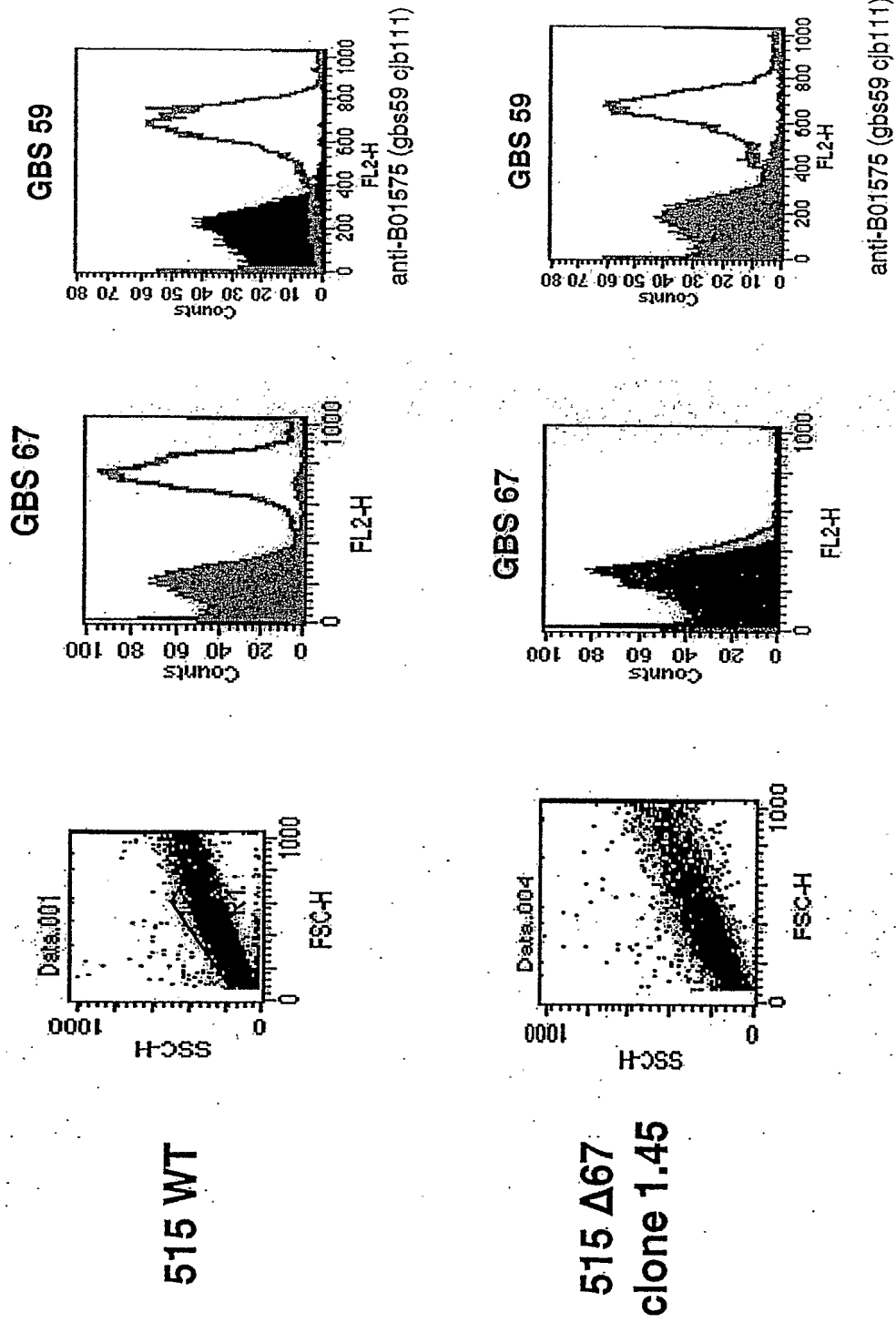


FIGURE 71

Complementation of GBS 515 KO 67 with pAM401-gbs80

GBS 80 forms a high molecular weight complex (pilus) in absence of GBS 67

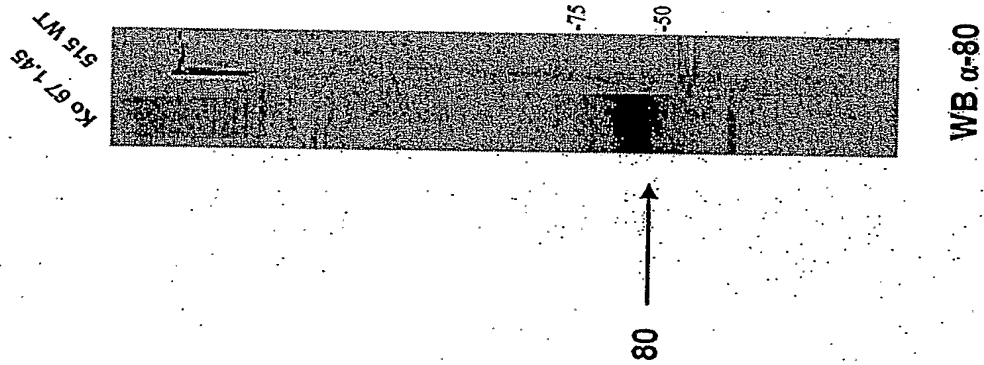


FIGURE 72

spyM6_0159 type 1 pilus present in M6

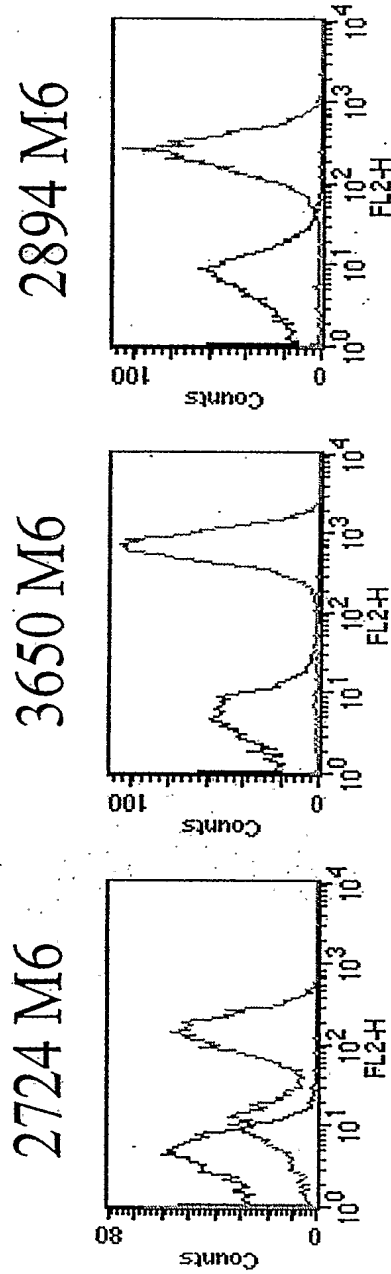


Figure 73

PCT/US05/27239 224/487

spyM6_0160 type 1 pilus present in M6

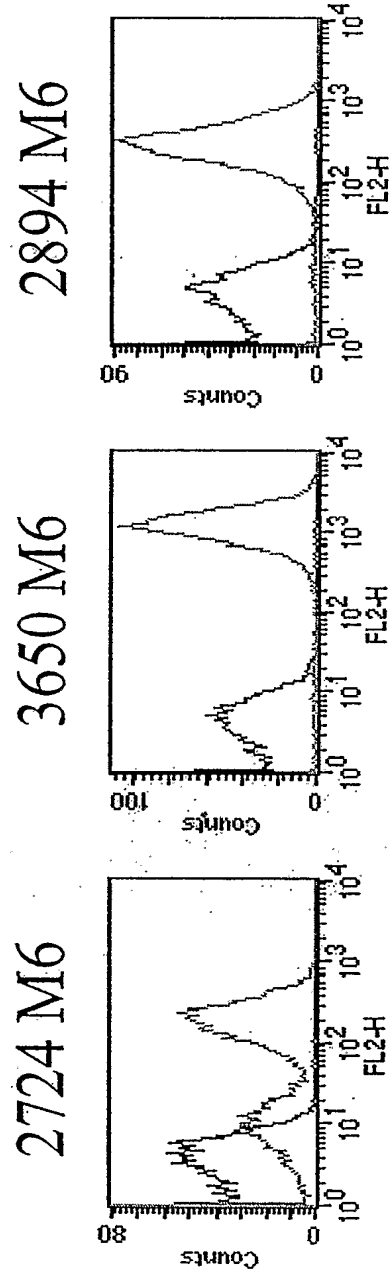


Figure 74

Gas15 type 2 pilus present in M1

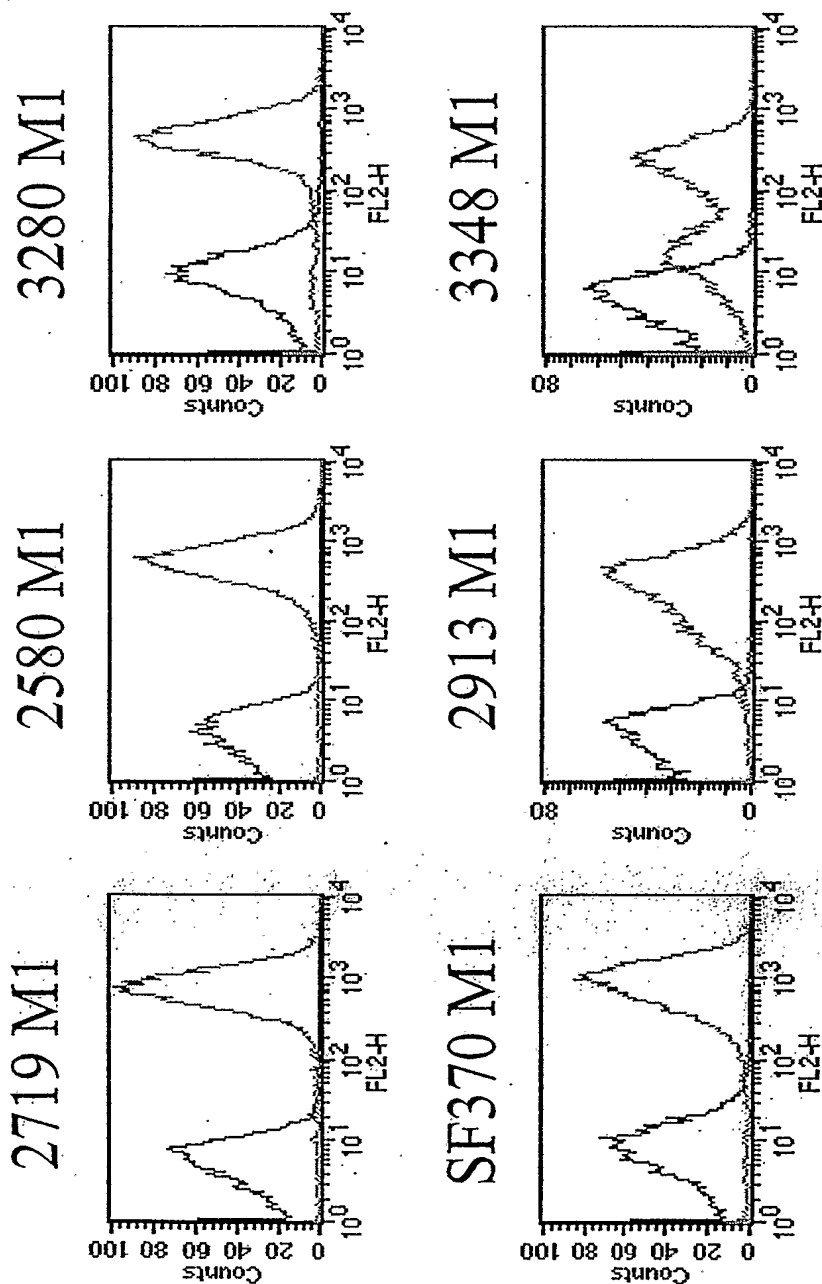


Figure 75

Gas16 type 2 pilus present in M1

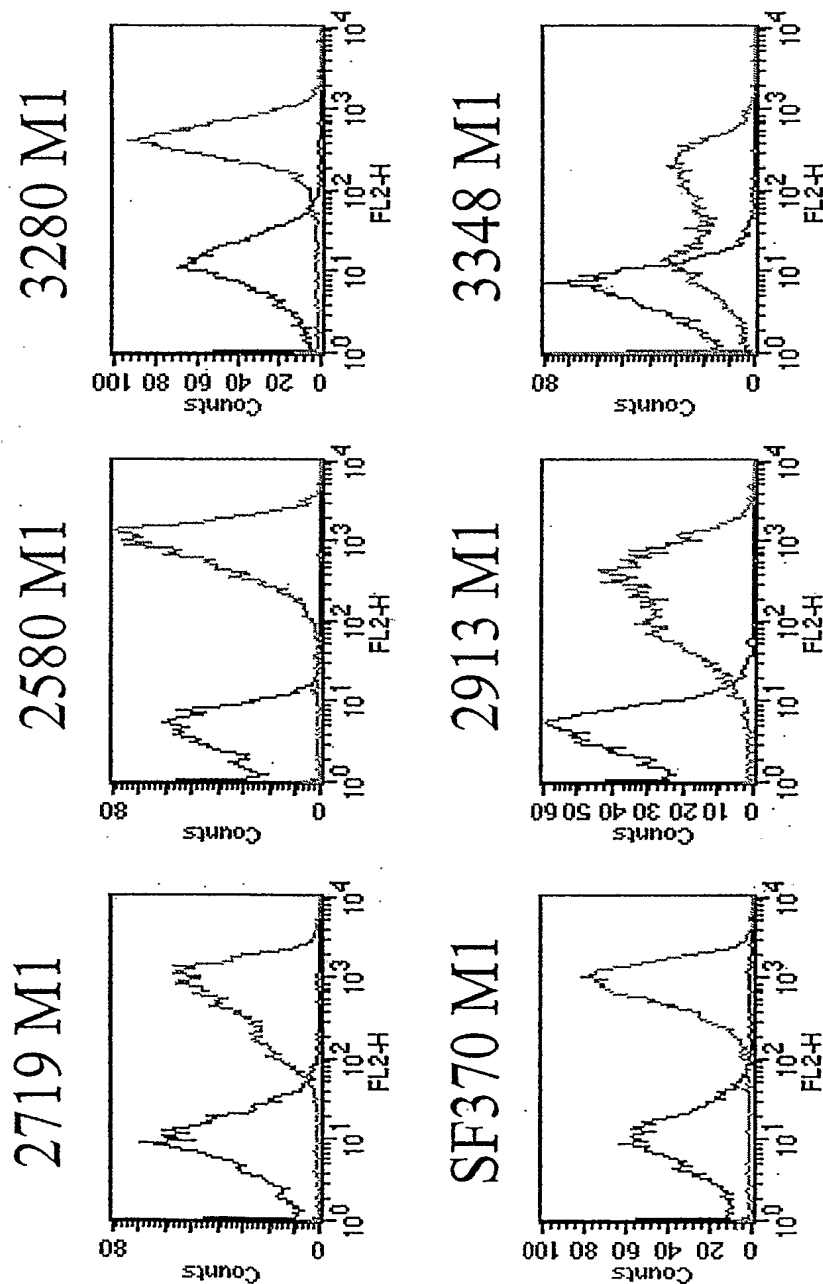


Figure 76

Gas18 serum 1 type 2 pilus present in M1

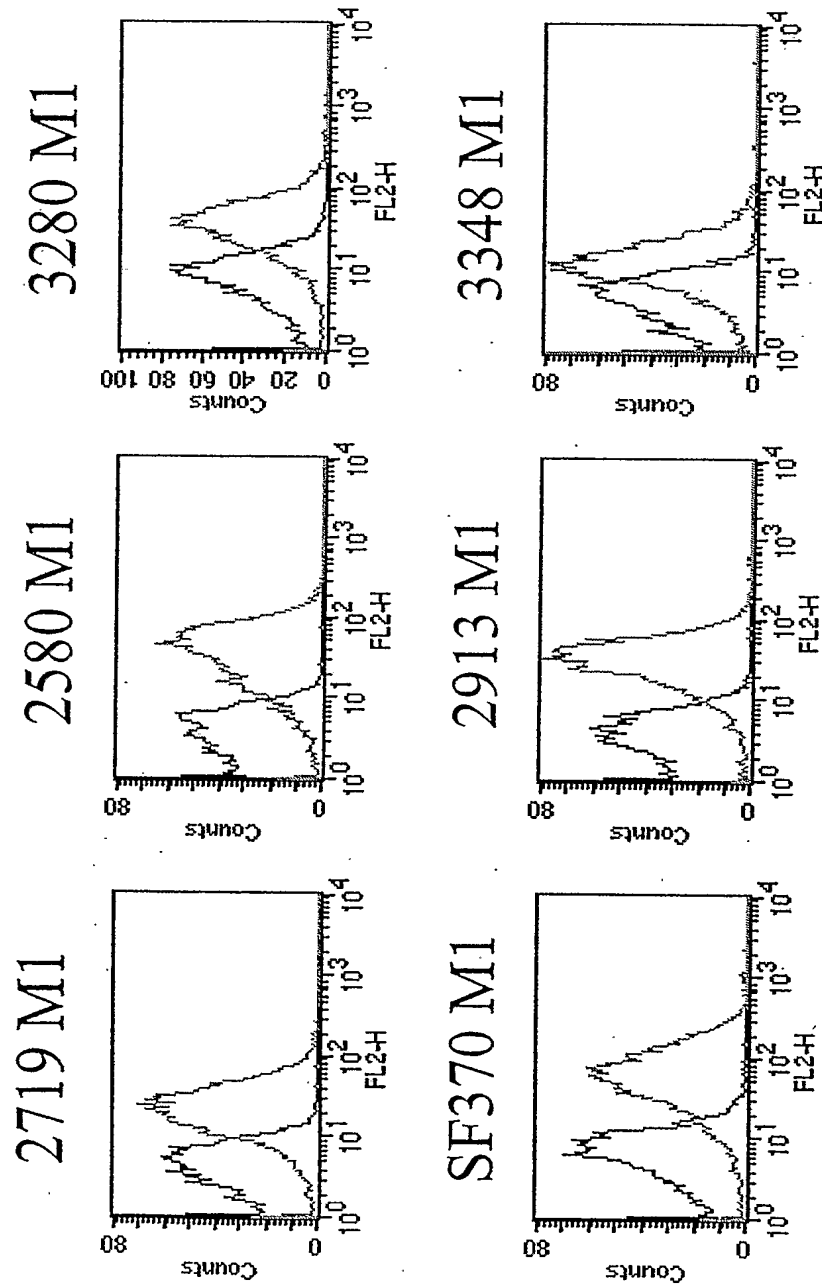


Figure 77

Gas18 serum 2 type 2 pilus present in M1

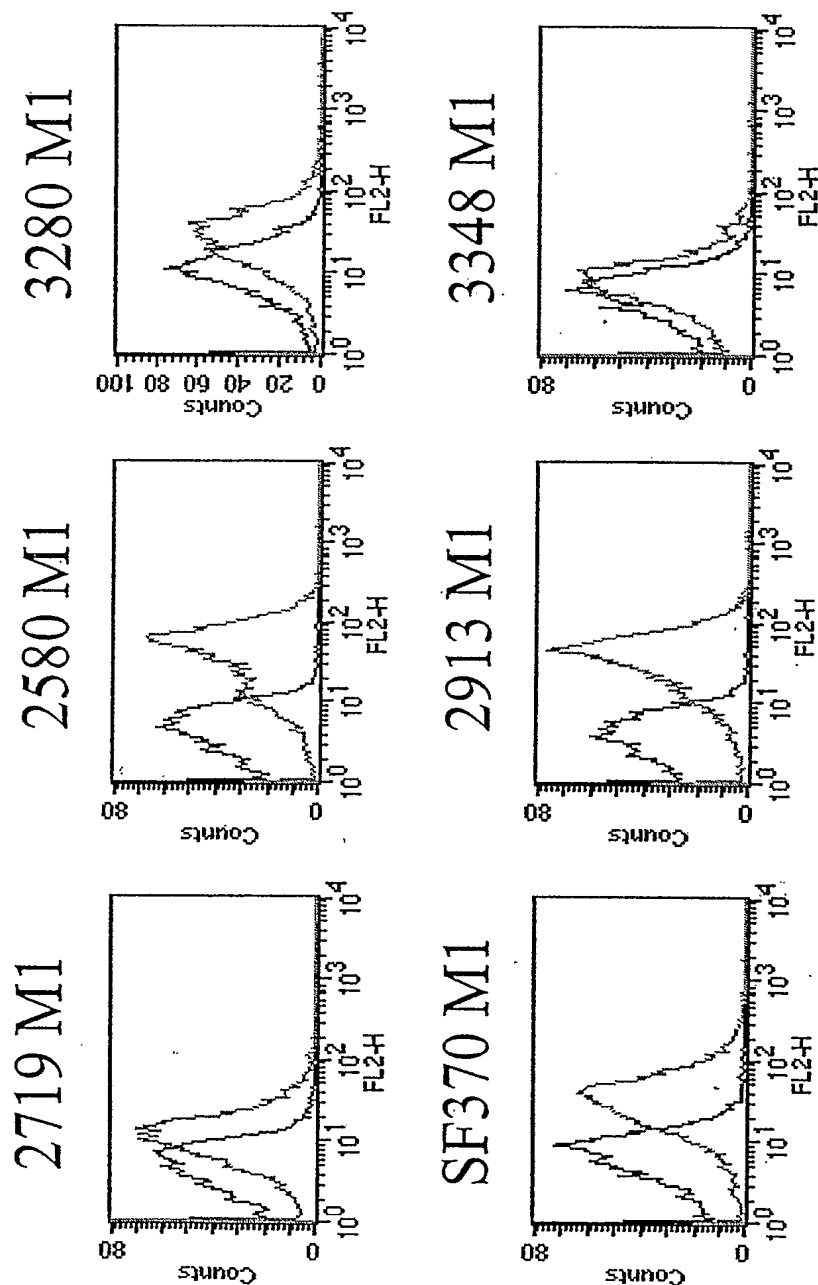


Figure 78

Gas16p2 type 2 pilus present in M1

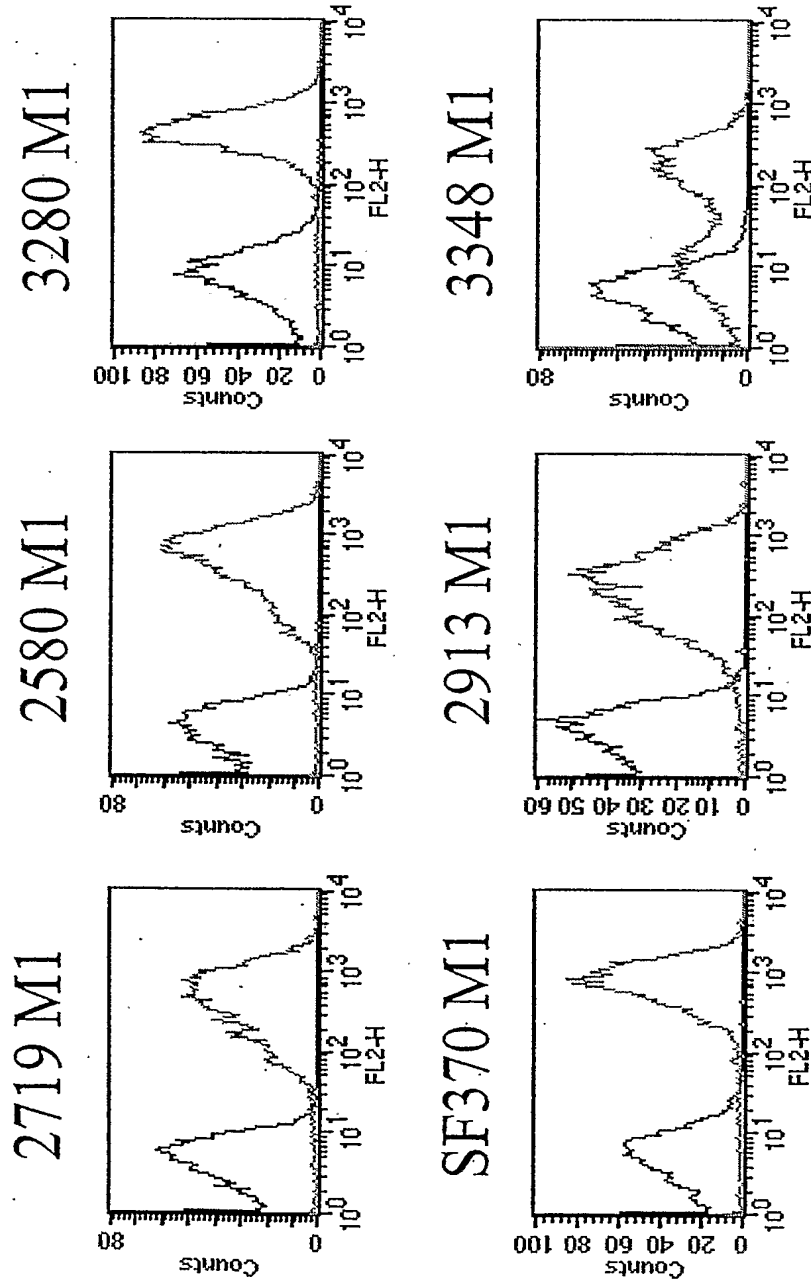


Figure 79

spyM3_0098 type 3 pilus present in M3

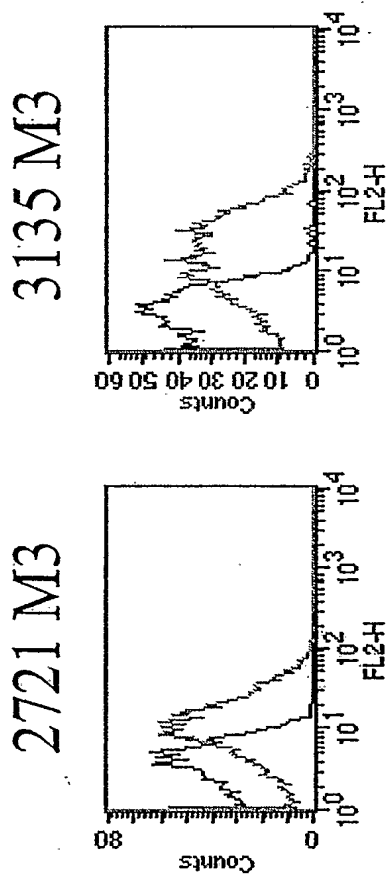
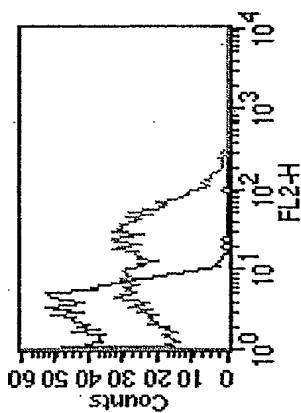


Figure 80

spyM3_0100 type 3 pilus present in M3

3135 M3



2721 M3

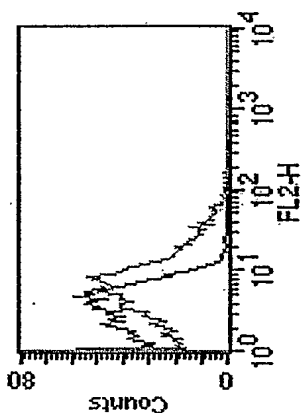


Figure 81

spyM3_0102 type 3 pilus present in M3

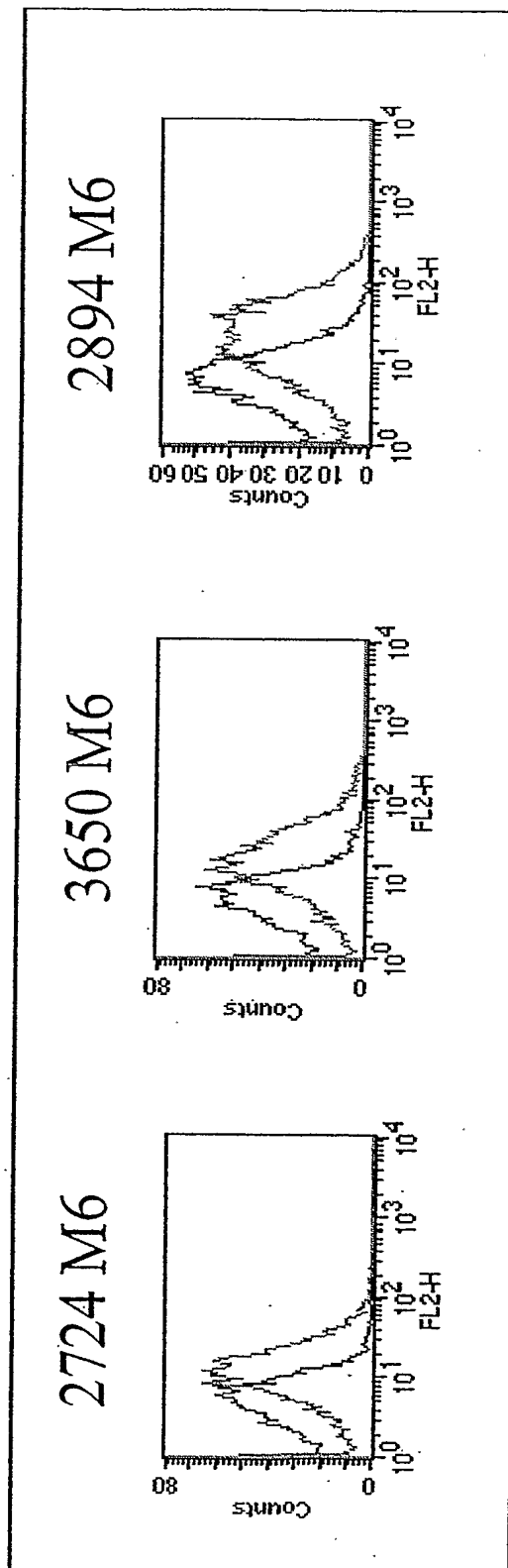
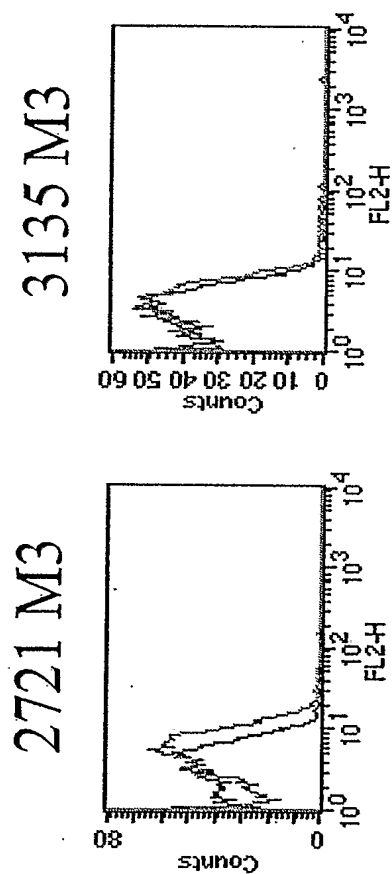
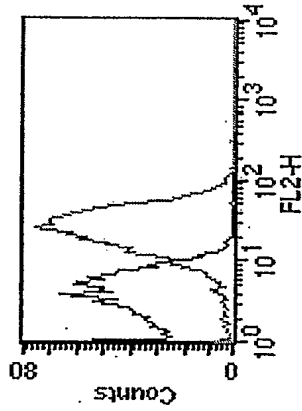


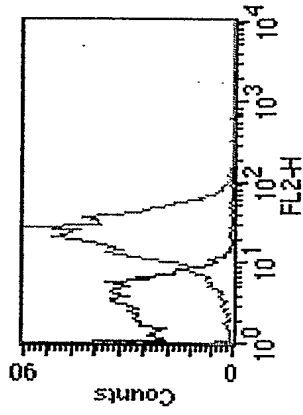
Figure 82

spyM3_0104 type 3 pilus present in M3

2721 M3



3135 M3



2728 M12

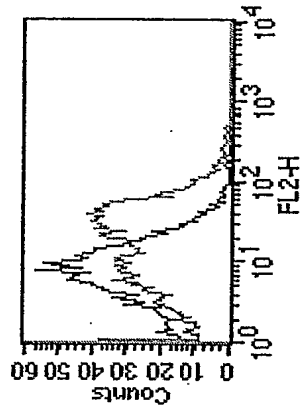
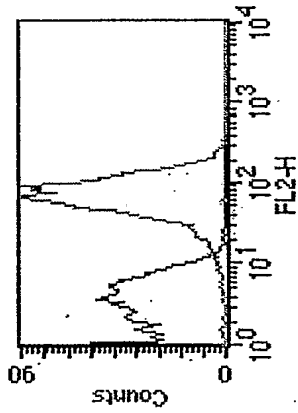


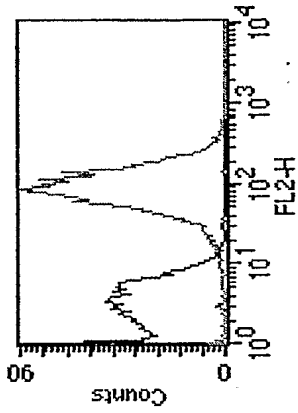
Figure 83

spyM3_0106 type 3 pilus present in M3

2721 M3



3135 M3



2728 M12

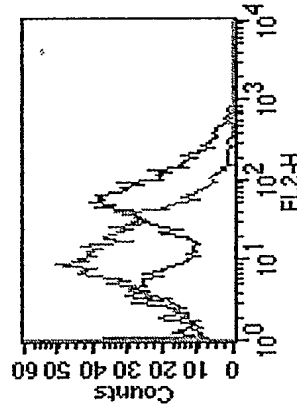
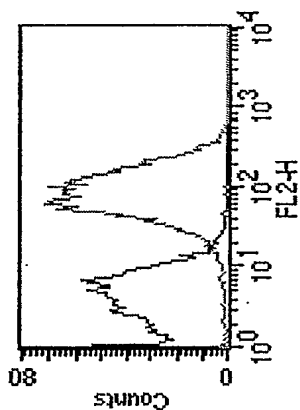


Figure 84

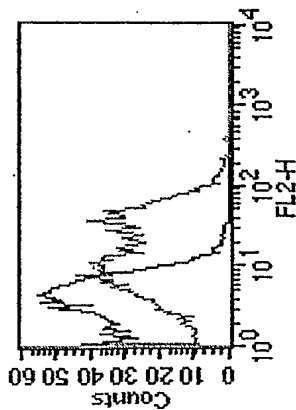
PCT/US05/27239 235/487

19224134 type 4 pilus present in M12

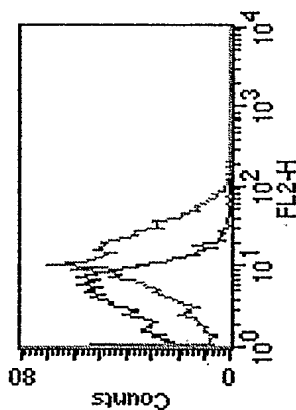
2728 M12



2724 M6



3650 M6



2894 M6

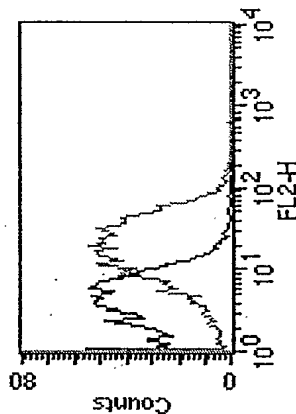


Figure 85

19224135 type 4 pilus present in M12

2728 M12

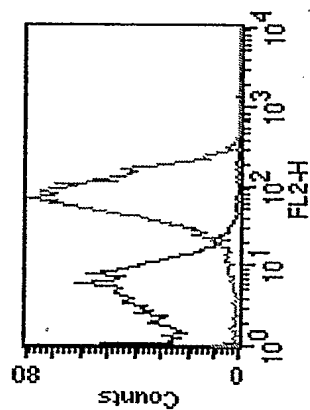


Figure 86

19224137 type 4 pilus present in M12

2728 M12

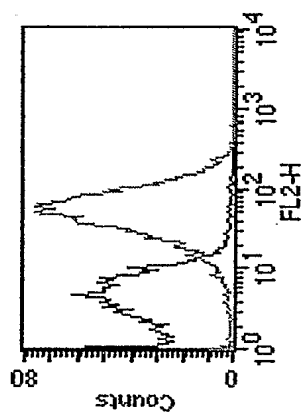


Figure 87

19224141 type 4 pilus present in M12

2728 M12

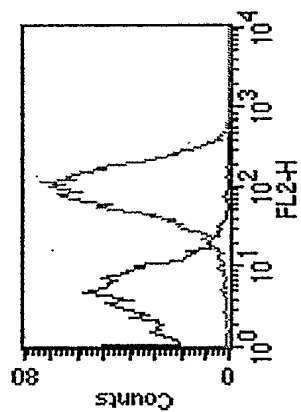
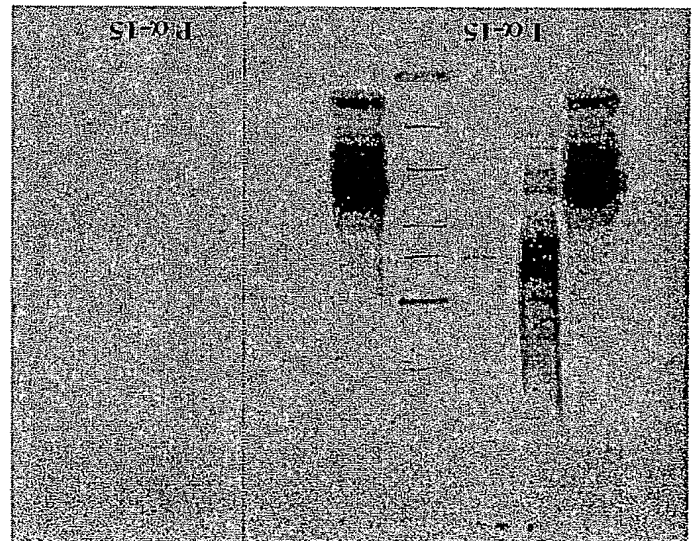


Figure 88

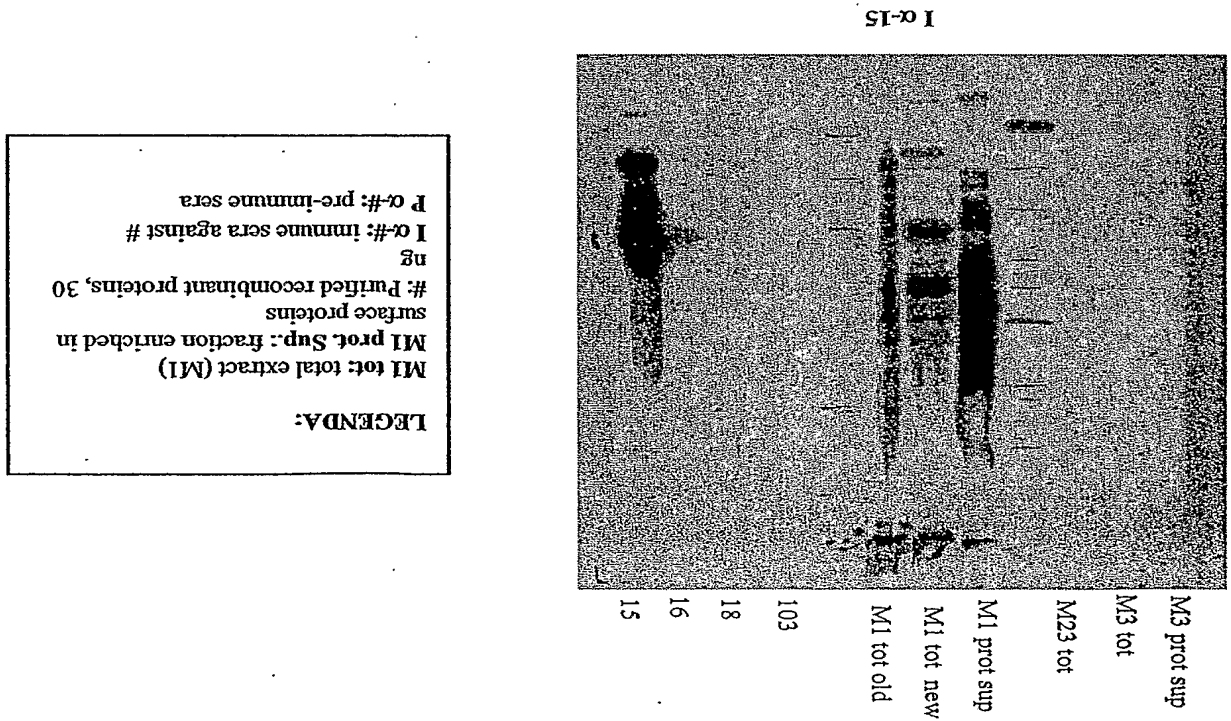
Figure 89

LEGEND:
 MI tot: total extract (MI)
 MI surf prot: fraction enriched in
 surface proteins
 #: Purified recombinant proteins, 30
 ng
 I α-#: immune sera against #
 P α-#: pre-immune sera



15
 MI surf prot
 MI tot
 15
 16
 18
 MI surf prot
 MI tot

Figure 90



PCT/US05/27239 341/487

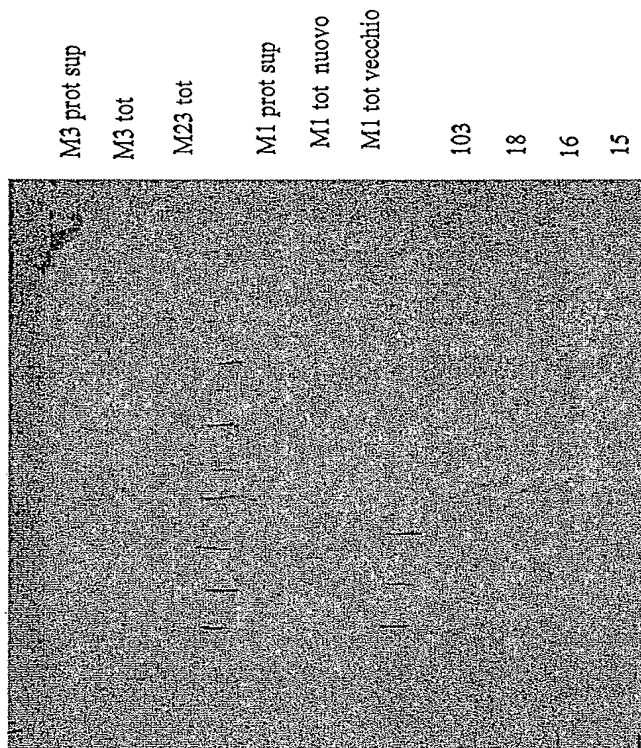
**LEGENDA:****M1 tot:** total extract (M1)**M1 prot. Sup.:** fraction enriched in surface proteins**#:** Purified recombinant proteins, 30 ng**I α-#:** immune sera against #**P α-#:** pre-immune sera**P α-15**

Figure 91

PCT/US05/27239

242/487

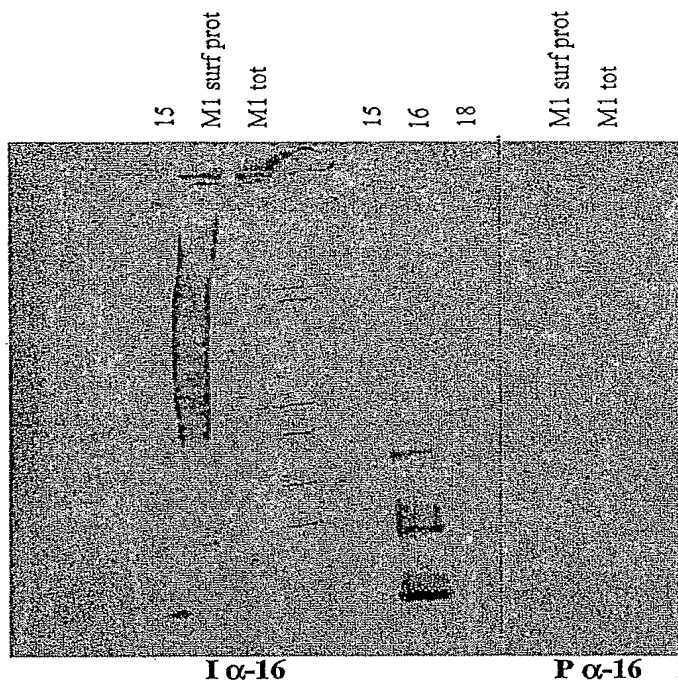
**LEGEND:****M1 tot:** total extract (M1)**M1 surf prot.:** fraction enriched in surface proteins**#:** Purified recombinant proteins, 30 ng**I α-#:** immune sera against #**P α-#:** pre-immune sera

Figure 92

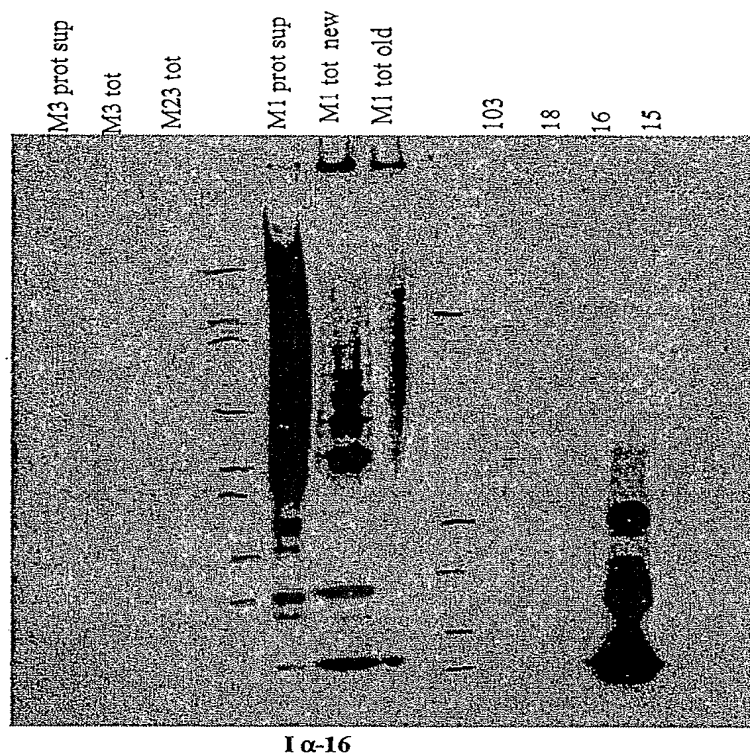
**LEGENDA:****M1 tot:** total extract (M1)**M1 prot. Sup.:** fraction enriched in surface proteins**#:** Purified recombinant proteins, 30 ng**I α -#:** immune sera against #**P α -#:** pre-immune sera

Figure 93

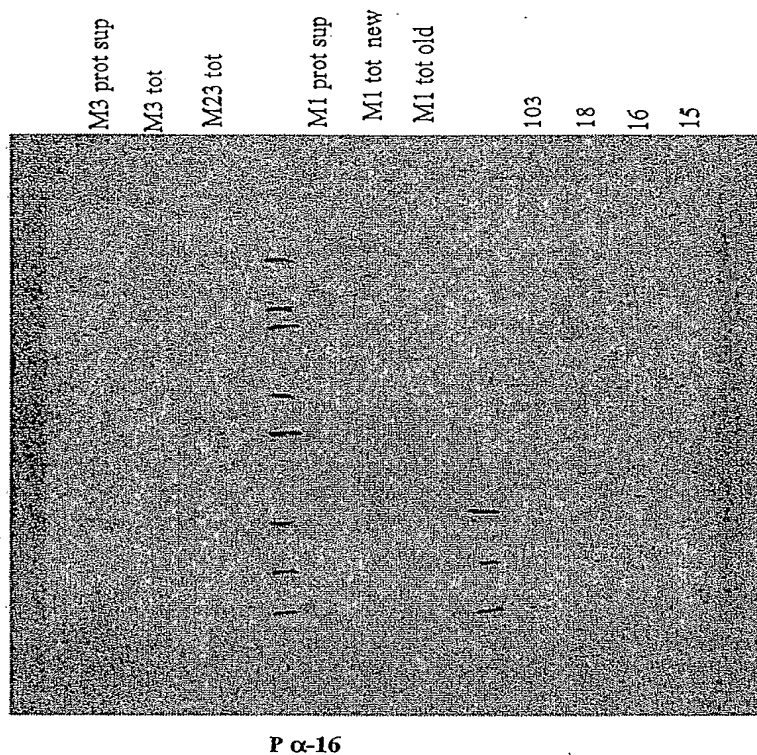
**LEGENDA:****M1 tot:** total extract (M1)**M1 prot. Sup.:** fraction enriched in surface proteins**#:** Purified recombinant proteins, 30 ng**I α-#:** immune sera against #**P α-#:** pre-immune sera

Figure 94

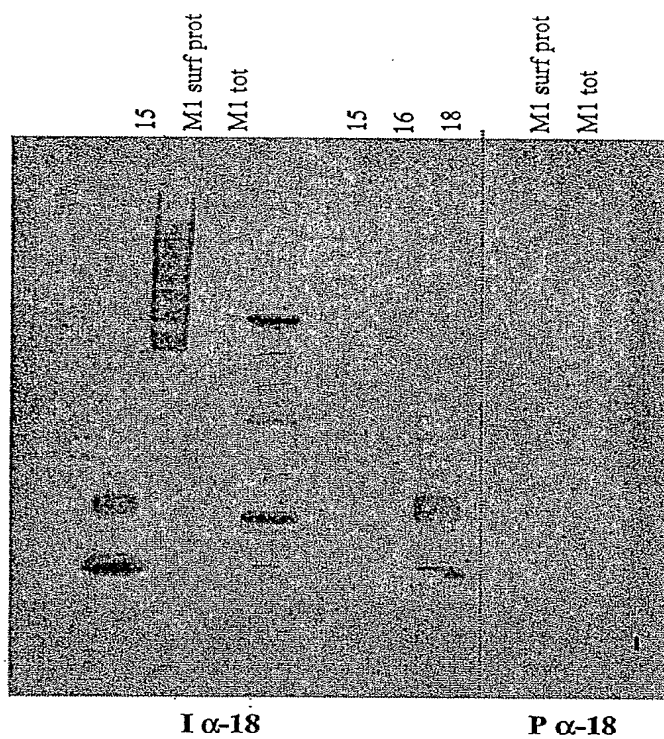
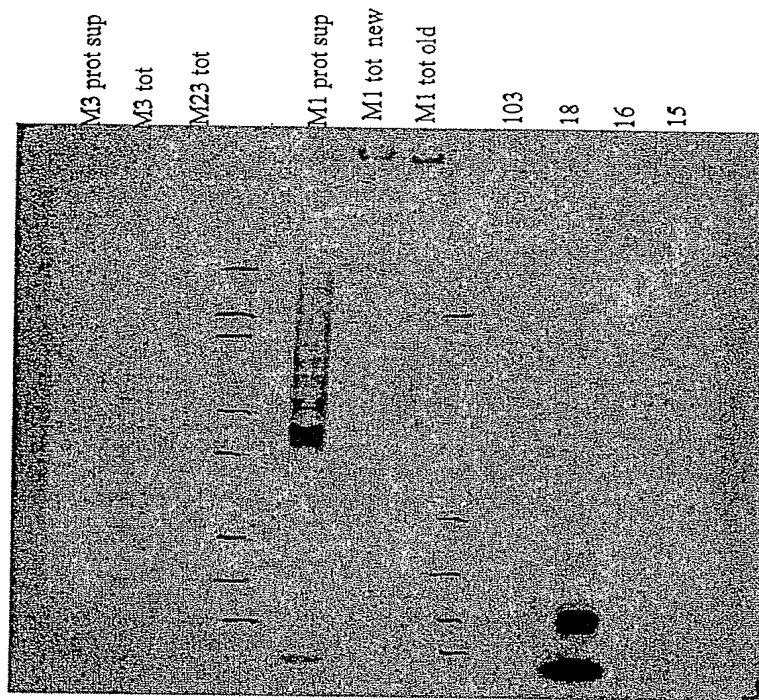
**LEGEND:****M1 tot:** total extract (M1)**M1 surf prot.:** fraction enriched in surface proteins**#:** Purified recombinant proteins, 30 ng**I α -#:** immune sera against #**P α -#:** pre-immune sera

Figure 95

PCT/US05/27239 246/487



I α-18

LEGENDA:

M1 tot: total extract (M1)
M1 prot. Sup.: fraction enriched in
surface proteins

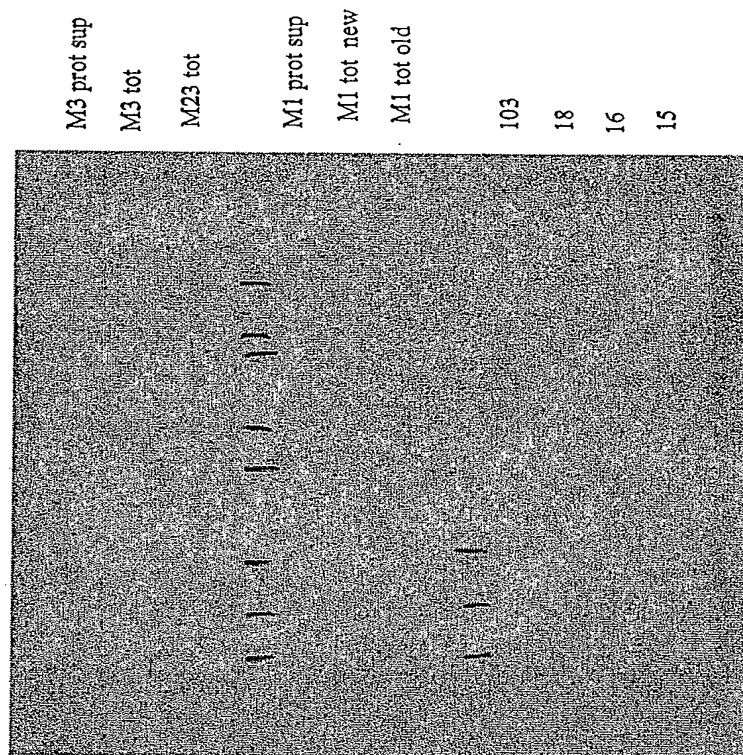
#: Purified recombinant proteins, 30
ng

I α-#: immune sera against #

P α-#: pre-immune sera

Figure 96

PCT/US2005/027239 247/487



LEGENDA:

M1 tot: total extract (M1)

M1 prot. Sup.: fraction enriched in surface proteins

#: Purified recombinant proteins, 30 ng

I α-#: immune sera against #

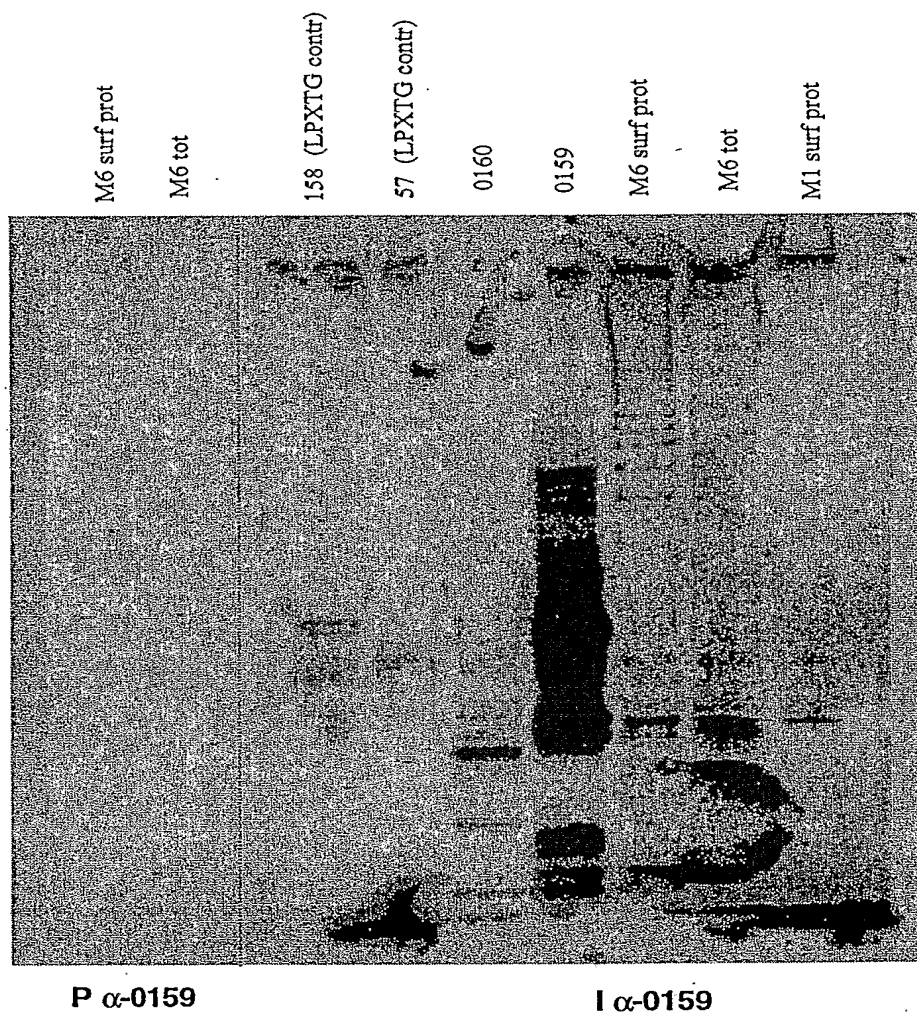
P α-#: pre-immune sera

P α-18

Figure 97

PCT/US05/27239 248/487

Figure 98

**LEGEND:****M6 tot:** total extract (M6)**M6 surf prot.:** fraction enriched in surface proteins**#:** Purified recombinant proteins, 30 ng**I α-#:** immune sera against #**P α-#:** pre-immune sera

PCT/US2005/027239 249/487

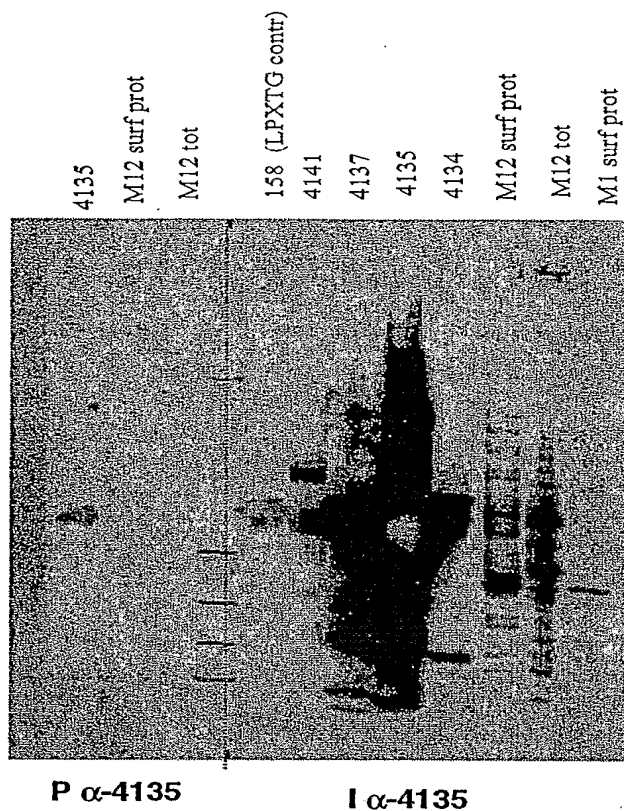
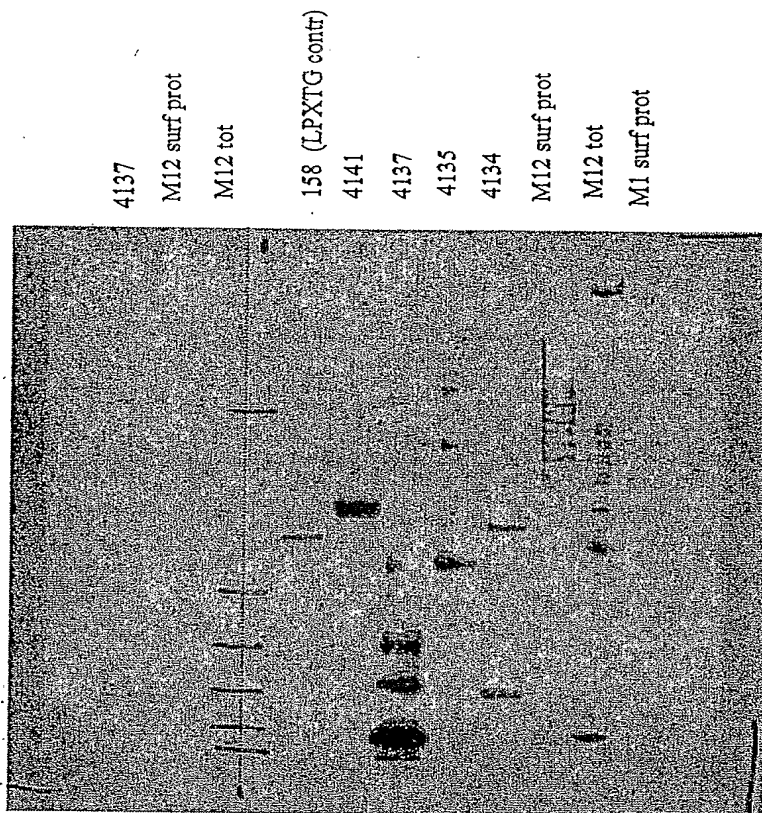
**LEGEND:****M12 tot:** total extract (M12)**M12 surf prot.:** fraction enriched in surface proteins**#:** Purified recombinant proteins, 30 ng**I α-#:** immune sera against #**P α-#:** pre-immune sera

Figure 99

PCT/US05/27239 250/487

**LEGEND:**

M12 tot: total extract (M12)

M12 surf prot.: fraction enriched in surface proteins

#: Purified recombinant proteins, 30 ng

I α-#: immune sera against #

P α-#: pre-immune sera

P α-4137

I α-4137

Figure 100

FIGURE 101

1 GACAGCTTCCCTTATACGACCGCTTCTATATCGGACCTTCCAGTTCTTGCTCTTTTA 60
CTGTTGGAAGGAATATGCTGGCGAAAGATATAGCCTGAGAAGTTCAAGAACGAGAAAAT
61 CCAGGACTAGCCGTTTCAGGTGCAACGATTGTGGTGGTTGTTAAATGGAACCAAGTCGT
GGTCCCTGATCGGCAAGTCCACGTTGCTAACAGCCACCAACAACAATTACCTTGGTCAGCA 120
121 TCAGTTGTGACAGAAATTTACCTTCTATCTTGGGATTCCTCGTTATGTTGGAGCTAGTGCC
AGTCAACACTGTCTTAAATGGAAGATAGAACCTTAAGGCAATACAAACCTCGATCACGG 180
181 TTAAGATTTTCAAAATTTGTGAAAGCCGGAGAACTCTTGAGCTTTGGGCAATGTTTTTG
AATTCTAAAAGTTTAAACACTTTCGGCCCTCTTGAGAACTCGAAACCCGTTAACAATAAC 240
241 CTCCTGGTCGCGATGGGAGTAGCTTTTGGCGTCAGATGGTGGCTATTCTGCTCTTGACC
GAGAACCAAGCGCTACCTTCATCGAAACAGCCAGTCGTACCAACCGATAAGCGAAGACTGG 300
301 AGCTATGTGAAAAACACGACTTCACCCCTTTTGGTAAATAACCGTATCGTGTGGTAGT
TCGATACACTTTTGTGCTGAAGTGGGAAAAAACCATTTATGGCATAGCACGAACCATCA 360
361 GTTTTGCTACTTTACAGTTTGTCCGTTTATTTGTATAAGAAAAACCTTGAGGGGTAAAC
CAAAACGATGAATGTCAAAACAGGCAATAAACAATATCTTTTGGAACTTCCCCCATTG 420
421 TCTTCAAGGTTTTTATACCTTAGAAATCTCTTCAAAACCGGCTCAGCTTTATCTGCAACC
AGAAGTTCAAAATATGAGAATCTTTTAGAGAAGTTTGGCGGCAATAGACCGTTGG 480

Figure 101A

481 TCATAAACAGTGTGTTTGAGCAGCCCTGCGGCTAGCTTCCTAGTTTCCTCTCTTCGATTTTCATT
-----+-----+-----+-----+-----+-----+-----+-----+
AGTTTGTGCACAAAACCTCGTCGGACGCGGATCGAAGGATCAACGAGAACTAAAGTAA 540

541 GAGCTTTAAATCCAGTCAGGGTAATCCCCAATAGCGGACACCTCTTTCTTTCTCTCGCTT
-----+-----+-----+-----+-----+-----+-----+-----+
CTCGAAATTTTAGGTCAGTCCCATTTAGGGTTATTCGCCCTGTGGAGAAAGAAAGAGCGAA 600

601 AATTCTTCATAGAGTTGCAGGGCTATTTGGCTTATCTGACTAGCATCTTGTGTTTTTGG
-----+-----+-----+-----+-----+-----+-----+-----+
TTAAGAAGTATCTCAACGTCGCCGATAAACCGAATAGACTGATCGTAGAACACAAAAACC 660

661 CAAGACTTTTTCGTTTGGTAAGAGTTGAAAAAGTCCCTCGTAGCGGATTTTCAAAATGACAA
-----+-----+-----+-----+-----+-----+-----+-----+
GTTCTGAAAAAGCAACCATTTCTCAACTTTTTCAGGACATCGCCTAAAGTTTTACTGTT 720

721 TTTTTCAGCTTTTTCCTTGTGATGTAGATTGAGAGGACTTTTCTGTATAGAAAGATCA
-----+-----+-----+-----+-----+-----+-----+-----+
AAAAAGTTCGAAAAAGAACAACTACATCTAACTCTCGCTGAAAAAGACTATCTTCTCAGT 780

781 GCTCTTTTGTGATATCTTCTCGGCACGGAGAAATCTTCCCGTAGGTTTTCCTTCCTGCCGA
-----+-----+-----+-----+-----+-----+-----+-----+
CGAGAAAAAACTATAGAGGAGCCGTGCTCTTTAGAGGGCATCCAAAAGAGGAACGGCT 840

841 TTGATTTACGATGCGATTGGATTGACTGGAGAGTTGTGAATGCCACGAGCCTTTTCGAT
-----+-----+-----+-----+-----+-----+-----+-----+
AACTAAATGCTACGCTAACCTAACTGACCTCTCAACACTTACGGTGTCTGGAAAGCTA 900

901 ACAGATCATAGCCTAGTCTACCAAAAACGGTCTATTAGGTTACTCAGGAACCTTCAAGTA
-----+-----+-----+-----+-----+-----+-----+-----+
TGTCATATCGGATCAGATGGTTTTTGGCAGATAATCCCAATGGAGTCCCTTGAAGTTTCA 960

253/487

Figure 101B

961 AATCAGCACCAAGTAAACAGCCCATTTGATGAAGACGTTCTACTGCTCTTTTCCCTACTC
-----+-----+-----+-----+-----+-----+-----+
TTAGTCGTGGTCATTTTTCGGGTAAACTACTTCTGCAAGATGACAGAAAAAAGGATGAG
-----+-----+-----+-----+-----+-----+-----+ 1020

1021 CATGAAATTTGGAATATCCATTTGTTTGAGAAAACTCTCAGCCCTGTTTCAGGTAGAATCA
-----+-----+-----+-----+-----+-----+-----+ 1080
GTACTTTAAACCTTTATAGGTAAACAAACTCTTTTAGGAGTCGGACAAGTCGATCTTAGT
-----+-----+-----+-----+-----+-----+-----+

1081 CTGTCAAAACCATGTGGTTTTTGATAATCACTCGCCATTTTAGCTAAGAAATTTGTGTAAAG
-----+-----+-----+-----+-----+-----+-----+ 1140
GACAGTTTGGTACACCAAAAACCTATTAGTGAGCGGTAAATCGATCTCTTAAACAACATTC
-----+-----+-----+-----+-----+-----+-----+

1141 AAACGCCCTGCGAAGCAGTTAGATGGAGTCTTTCAGATATCTTTTGAATGAGGCGAG
-----+-----+-----+-----+-----+-----+-----+ 1200
TTTCCGGACGCCCTTCGTCATCTACCTCAAGAAAGGCTATAGAAAAAACTTACTCCGCTC
-----+-----+-----+-----+-----+-----+-----+

1201 CAATTTGACCGCTGACTTGATACCGAGTTATTTCTGTCTACATCCAAATAGGCTTCGT
-----+-----+-----+-----+-----+-----+-----+ 1260
GTTAAAACTGGCGACTGAACATATGGCTCAAAATAAAGACAGTGTAGTTTATCCGAAGCA
-----+-----+-----+-----+-----+-----+-----+

1261 CAATGCTCATGGGTTCAATCAAAATCTGTATAGCGCTTAAATAATAGCTCGAATCCGGAGTC
-----+-----+-----+-----+-----+-----+-----+ 1320
GTTACGAGTACCCAAAGTTAGTTTAGACATATCCGGAATTTTATCGAGCTTAGGCCCTCAG
-----+-----+-----+-----+-----+-----+-----+

1321 CCACAGACTTGTATTTCTCATAATTCCTGAGATAAAGACAGCCCTGGGACACACGTTTCAT
-----+-----+-----+-----+-----+-----+-----+ 1380
GGTGTCTGAACATAAAGAGTATTAAGGACTCTATTCTGTGCGGACCCCTGTGTGCAAGTA
-----+-----+-----+-----+-----+-----+-----+

1381 AAGCTTCCTTGGAACTCATGGCAGAAATGGACACCAAAAGCTCTTGCCTCATTAACACAGG
-----+-----+-----+-----+-----+-----+-----+ 1440
TTCCGAAGGAACCTTGAGTACCGTCTTACCTGTGGTTTTTCGAGAAGGAGTATGTGATGCC
-----+-----+-----+-----+-----+-----+-----+

PCT/US05/27239
254/487

Figure 101C

1441 TAGAACGACTCCCGCTCCACCTGTTTGGCGAGGTCGCTTCCAATAATGACAGGTTTTC
-----+-----+-----+-----+-----+-----+-----+
1500 ATCTTTGCTGAGGGCAGGTGGACAAACGGCTCCAGCGAGGTTATTTACTGTCCAAAG
-----+-----+-----+-----+-----+-----+-----+
1501 CTCTGAGTTTAGGATTATCCCTGATTCCACTGCAGCAAAAAGGCATCCATGCAATAT
-----+-----+-----+-----+-----+-----+-----+
1560 GAGACTCAAATCCTAATAGGGACTAAAGGTGACGTCGTTTTTTTCCGTAGGTACAGTTATA
-----+-----+-----+-----+-----+-----+-----+
1561 GGATGATTTTCTCTTGACAAATCATTTAACAAAGGAAAAATCAACATGCCCTAGCACCTTTT
-----+-----+-----+-----+-----+-----+-----+
1620 CCTACTAAAAGAACTGTTTAGTAAATTGTTTCCCTTTTAGTTGTACGGATCGTGGAAAA
-----+-----+-----+-----+-----+-----+-----+
1621 TATACTCTTCGAAAAATCTCTTCAAAACCGTCAGCTTCCATCTGCAACCTCAAAACAGTA
-----+-----+-----+-----+-----+-----+-----+
1680 ATATGAGAAGCTTTTAGAGAAGTTTGGTCAGTCGAGGTAGACGTTGGAGTTTGTGCTAT
-----+-----+-----+-----+-----+-----+-----+
1681 TTTTGAGCTGACTCGTCAGTTCTATTATACAACTCAAGCAGTCGTTTGACGAGCCTGC
-----+-----+-----+-----+-----+-----+-----+
1740 AAAACTCGACTGAAGCAGTCAAGATAAATGTTGGAGTTTCGTACAGAACTCGTCGGAGG
-----+-----+-----+-----+-----+-----+-----+
1741 GGCTAGTTTCTCTAGTTTGGCTTTTTCGATTTCCTCATTTGAGTGTAACCTGCTTATTTTCTTTTAT
-----+-----+-----+-----+-----+-----+-----+
1800 CCGATCAAAGGATCAAACGAAAAGCTAAAGGTAACACATTCAGCAATAAAGAAAAATA
-----+-----+-----+-----+-----+-----+-----+
1801 TATACCTTTTTCCTGAAAAAAGAAAAAGGACTTTATTTTTCATAAATAATATATACA
-----+-----+-----+-----+-----+-----+-----+
1860 ATATGGGAAAAAAGACTTTTTCCTTTTCTCTGAAATAAAAAAAGTTTATATATATATGT
-----+-----+-----+-----+-----+-----+-----+
1861 GTTTGAAATAAATAATAGACTGTTTGTAGAAAAAGTGTAAAAATAGGAATTTTTCACCT
-----+-----+-----+-----+-----+-----+-----+
1920 CAAACTTTATTTTATATCTGACAAATCTTTTCTTTTCTTTTTCACATTTTATTCCTTAAAAAGTGA
-----+-----+-----+-----+-----+-----+-----+

Figure 101D

1921 TGTTGAAATCGGTTACTTTATGGTATACTTGTCTCANGAATGTAAACAGATGACTGTTACT + 1980
-----+-----+-----+-----+-----+-----+-----+-----+
ACAACTTTAGCCCAATGAATACCATATGAACAGAGTACTTACATTGTCTACTGACAAATGA
-----+-----+-----+-----+-----+-----+-----+-----+
AGAAAAAGAGGACATTAATATGTTGTTAAGACAGTTGTTGAAGCACAAAGATATTTTTC + 2040
-----+-----+-----+-----+-----+-----+-----+-----+
TCCTTTTTCCTCTGTAATTATACCAACAATCTGTCAACAACACTGCTGTTCTATATAAAAC
-----+-----+-----+-----+-----+-----+-----+-----+
c M V V K T V V E A Q D I F D - this orf is the homologue of sp0459, a formate
acetyltransferase (pfl). It is out of the
pilus locus
2041 ACAAGCTTGGGAAGGCTTCAAGCGTAGATGGAAGAAAAAGCAAGTATCACGCT + 2100
-----+-----+-----+-----+-----+-----+-----+-----+
TGTTTCGAACCCCTTCCGAAGTTTCCGCATCTAACCTTCTTTTTCGTTTCACATAGTGCGA
-----+-----+-----+-----+-----+-----+-----+-----+
c K A W E G F K G V D W K E K A S V S R F -
2101 TTGTACAAGCTAACTACACACCTTATGATGGAGACGAAAGCTTCTTGCAGGACCAACAG + 2160
-----+-----+-----+-----+-----+-----+-----+-----+
AACATGTTGATGATGTGGAATACTACCTCTGCTTTCGAGGAACGCTCGTTGTC
-----+-----+-----+-----+-----+-----+-----+-----+
c V Q A N Y T P Y D G D E S F L A G P T E -
2161 ACGGTTCACTTCACATCAGAAAAATTGTAGAAGAACTAAAGCACACTAGCAAGAAACTC + 2220
-----+-----+-----+-----+-----+-----+-----+-----+
TCGCAAGTGAAGTGTAGTCTTTTAAACATCTTCTTTGATTTCTGCTGTGATGCTTCTTTGAG
-----+-----+-----+-----+-----+-----+-----+-----+
c R S L H I K K I V E E T K A H Y E E T R -
2221 GTTTCCTCAATGGACACTGCTCCAACATCTATCGCTGATATCCCTGCTGGATTATCGACA + 2280
-----+-----+-----+-----+-----+-----+-----+-----+
CAAAGGTTACCTGTGACAGGTTGTAGATAGCGACTATAGGACGACCTAAATAGCTGT
-----+-----+-----+-----+-----+-----+-----+-----+
c F P M D T R P T S I A D I P A G F I D K -
2281 AAGAAAATGAAGTTATCTTTGGTATCCAAAATGATGAACCTCTTCAAAATGGAACCTTCATGC + 2340
-----+-----+-----+-----+-----+-----+-----+-----+
TTCCTTTACTTCAATAGAAACCATAGGTTTACTACTTGAAGATTAACTTGAAGTACG

Figure 101E

C E N E V I F G I Q N D E L F K L N F M P -
2341 CAAAAGGTGGTATCCGATGGCTGAAACTACTTTAAAGAAATGGATACGAACACGACC
-----+-----+ 2400
GTTTTCCACCATAGGCATACCGACTTTGATGAAATTTCTTTTACCTATGCTTGGTCTGG
-----+-----+
C K G G I R M A E T T L K E N G Y E P D P -
2401 CAGCTGTTACGAAATCTTCACTAAATATGTAACAACAGTTAACGACGGTATTTCCGCTG
-----+-----+ 2450
GTCGACAAGTGTCTTTAGAAAGTGATTTATACATTTCTGTGTCATTGCTGCCATTAAGGCAC
-----+-----+
C A V H E I F T K Y V T T V N D G I F R A -
2461 CCTACACTCAAAATATTCGTCGGCTCGTCATGACACACACTGTAACCTGGTCTTCCAGATG
-----+-----+ 2520
GGATGTGAAGTTTATAACGACGCCGACAGTACCTGTGTGTGACATTGACCAGAAAGGCTTAC
-----+-----+
C Y T S N I R R A R H A H T V T G L P D A -
2521 CATACTACCGGACGATATCATCGGTGTTTACGACGCTTCTCTCTTTACGGTGCAGACT
-----+-----+ 2580
GTATGAGTGGCCTGCATAGTAGCCACAATGCGTGCAGAACGAGAAATGCCACGCTCTGA
-----+-----+
C Y S R G R I I G V Y A R L A L Y G A D Y -
2581 ACTTGATGCAAGAAAAGTAAACGACTGGAATGCAATCAAGAAATCGATGAAGAAACAA
-----+-----+ 2640
TGAACACTGCTCTTTTTCATTGCTGACCTTACGTTAGTTCTTTTAGCTACTTCTTTGTT
-----+-----+
C L M Q E K V N D W N A I K E I D E E T I -
2641 TCCGCTCTCGTGAAGAAGTAAACCTTCAATACCAAGCATTCGAACAAGTTGTTCCGCTGG
-----+-----+ 2700
AGGCAGAACCACTCTTTCATTGGAAGTTATGTTCTGTAACGTTGTTCAACAAGCGGACC
-----+-----+
C R L R E E V N L Q Y Q A L Q Q V V R L G -
2701 GTGACCTTTACGGGTTGATGTTCCGCAACCCAGGATGAACGTGAAGAGCAATCCAAAT
-----+-----+ 2760

257/487

Figure 101F

CACTGGAAATGCCCAACTACAAGCGTTTGGTCGCTACTTGCACCTTCTTCCTTAGGTTA
D L Y G V D V R K P A M N V K E A I Q W -
2761
GGGTTAACATTGCTTTTCATGGCTGTCTGCGGTGTGATTAACGGTGTCTACATCTCTAG
-----+-----+-----+-----+-----+-----+-----+-----+ 2820
CCCAATTGTAACGAAAGTACCGACAGACGGCACACACTAATGCCACACGATGTAGAGATC
V N I A F M A V C R V I N G A A T S L G -
2881
GTCTGTACCAATCGTATTGGACATCTTTGACAGAACGTGACCTTCTCGTGGTACATTTA
-----+-----+-----+-----+-----+-----+-----+-----+ 2880
CAGCACATGGTAGCATAACCTGTAGAAAGCTCTTGCTGCTGGAACGAGCACCATGTAAAT
R V P I V L D I F A E R D L A R G T F T -
2941
CTGAATCAGAAATCCAAGAAATTCGTTGATGATTTGCTTATGAAACTTCGTACAGTTAAAT
-----+-----+-----+-----+-----+-----+-----+-----+ 2940
GACTTAGTCTTTAGGTTCTTAAGCAACTACTAAAGCAATACTTTGAAGCATCTCAATTTA
E S E I Q E F V D D F V M K L R T V K F -
3001
TTCTCGTACCAAGCTTATGACCAATTGTACTACTAGTGACCCACCTTTATCACAACCT
-----+-----+-----+-----+-----+-----+-----+-----+ 3000
AAGAGCATGGTTTCGAATACTGGTTAACATGAGTCCACTGGGTGGAAATAGTGTGAA
A R T K A Y D Q L Y S G D P T F I T T S -
3061
CTATGGCTGGTATGGGTAAACGACGGTCTGTCACCGTGTACTAAGATGGACTACCGTTTCT
-----+-----+-----+-----+-----+-----+-----+-----+ 3060
GATACCGACCATACCCATTTGCTGCCAGCTGCGCACATGATCTCTACCTGATGGCAAGA
M A G M G N D G R H R V T K M D Y R F L -
3121
TGACACTCTTGACAACTCGGTAACTCACCAGAACCAAACTTACAGTTCTTTGGACTG
-----+-----+-----+-----+-----+-----+-----+-----+ 3120
ACTGTGAGAACTGTTGTAGCCATTGAGTGGTCTTGGTTTGAAGTGTCAAGAAACCTGAC
N T L D N I G N S P E P N L T V L W T D -

PCT/US05/27239 258/487

Figure 101G

3121 ~ACAAATGCCATACAACTCCGTCGCTACTGTATGCACATGAGCCACAAACACTCTCTTA
+-----+-----+-----+-----+-----+-----+ 3180
TGTTTAAGCGTATGTTGAAGGACGCGATGACATACGTGTACTCGGTGTTGTGAGAAGAT
+-----+-----+-----+-----+-----+-----+
C K L P Y N F R R Y C M H M S H K H S S I -
+-----+-----+-----+-----+-----+-----+
3181 TCCAAATCGAAGGTGTAAACAACAATGGCTAAAGACGGATATGGTGAATGAGCTGTATCT
+-----+-----+-----+-----+-----+-----+ 3240
AGGTTATGCTTCCACATGTTGTGTACCGATTCTCTGCCTATACCACCTTACTCGACATAGA
+-----+-----+-----+-----+-----+-----+
C Q Y E G V T T M A K D G Y G E M S C I S -
+-----+-----+-----+-----+-----+-----+
3241 CATGCTGTGTCTCCACTTGATCCAGAAAATGAAGAACAACGCCACACATCCAGTACT
+-----+-----+-----+-----+-----+-----+ 3300
GTACGACACACAGAGGTGAACGTAGGCTCTTTACTTCTTGTGCGGTGTGTAGGTCATGA
+-----+-----+-----+-----+-----+-----+
C C V S P L D P E N E Q R H N I Q Y F -
+-----+-----+-----+-----+-----+-----+
3301 TCGGTCTCGTGTAAACGTTCTTAAAGCCCTTCTTACTGTTTGAATGGTGGTTACGACG
+-----+-----+-----+-----+-----+-----+ 3360
AGCCACGAGCACATTTGCCAAGATTTCGGGAAGAAATGACCAAACTTACCACCAATGCTGC
+-----+-----+-----+-----+-----+-----+
C G A R V N V L K A L L T G L N G G Y D D -
+-----+-----+-----+-----+-----+-----+
3361 ATGTTACAAAGACTACAAAGTATTTGATATCGAACCATCCGTGACGAAGTTCTTGAAT
+-----+-----+-----+-----+-----+-----+ 3420
TACAGTGTCTCTGATGTTTCTATAAATACTATAGCTTGGTTAGGCACCTGCTTCAACAACCTTA
+-----+-----+-----+-----+-----+-----+
C V H K D Y K V F D I E P I R D E V L E F -
+-----+-----+-----+-----+-----+-----+
3421 TTGAATCAGTTAAAGCGAACTTTGAAAATCTCTTGTACTGGTTGACTGACACTTACGTAG
+-----+-----+-----+-----+-----+-----+ 3480
AACTTAGTCAATTCGCTTGAACCTTTTATAGAGAACTGACCAACTGACTGTGAATGCATC
+-----+-----+-----+-----+-----+-----+
C E S V K A N F E K S L D W L T D T Y V D -
+-----+-----+-----+-----+-----+-----+
3481 ATGCTTGAACATCATCCACTACATGATAGGTACAACTACGAAGCTGTTCAAATGG
+-----+-----+-----+-----+-----+-----+ 3540
TACGGAACCTTGTAGTAGGTGATGTACTGACTATCCATGTTGATGCTTCGACAAGTTTACC
+-----+-----+-----+-----+-----+-----+

PCT/US05/27239 259/487

Figure 101H

C A L N I I H Y M T D R Y N Y E A V Q M A -
3541 CCTTCTGCGCAACTAAACAACGTCGCCAACATGGGATTCGGTATCTGTGGATTGCTTAACA
+-----+ 3600
GGAAGACGGTTGATTTGTTGACGGTTGTACCCCTAAGCCATAGACACCTAAACGATTGT
+-----+
C F L P T K Q R A N M G F G I C G F A N T -
3601 CTGTTGATACATTGTCAGCTATCAAAATACGCTACAGTTAAACCAATCCGTGACGAAGATG
+-----+ 3660
GACAACTATGTAACAGTCGATAGTTTATGCGATGTCAATTTGGTTAGGCACCTGCTTCTAC
+-----+
C V D T L S A I K Y A T V K P I R D E D G -
3661 GCTACATCTACGATTACGAACAATCGGTGACTACCCACGCTGGGGTGAAGATGACCCAC
+-----+ 3720
CGATGATGCTAATGCTTTGTTAGCCACTGATGGTGGCGACCCACTTCTACTGGGTG
+-----+
C Y I Y D Y E T I G D Y P R W G E D D P R -
3721 GTTCAACGAATTGGCAGAAATGGTTGATCGAAGCTTACACAACCTCTCTACGTAGCCACA
+-----+ 3780
CAAGTTGCTTAACCGTCTTTACCACTAGCTTCGAATGTGTTGAGCAGATGCATCGGTGT
+-----+
C S N E L A E W L I E A Y T T R L R S H K -
3781 AACTATACAAAGACGAGAGCTACAGTATCCTTTTACAAATCAGATCTAAGTTGCTT
+-----+ 3840
TTGATATGTTTCTGCGTCTTCGATGTCATAGTGAAACTGTTAGTTAGATTCCAAACGAA
+-----+
C L Y K D A E A T V S L L T I T S N V A Y -
3841 ACTCTAAACAACTGGTAACACCAAGTTCACAAAAGTGTATACCTCAACGAAGATGGTT
+-----+ 3900
TGAGATTGTTTACCATTTGAGTGGTCAAGTGTTCACATATGGAGTTGCTTCTTACCAA
+-----+
C S K Q T G N S P V H K G V Y L N E D G S -
3901 CTGTGAACCTTGCTAACTTGAATCTTCTCACCAGTGTCAACCATCTTAAACAAAGCTA
+-----+ 3960

PCT/US05/27239 260/487

Figure 101I

GACACTTGAACAGATTGAACTTAAAGAGTGTGTCACGATGGGTAGATTGTTTCGAT
V N L S K L E F S P G A N P S N K A K -
AAGTGGTGGTTCGAAAACCTTGAACCTCATTCTAGCCCTGACTTTAGTTATGCACCTG + 3961
TTCCACCAACCAACGTTTGAACCTTCAGTGAAGATCGGAACCTGAAATCAATACGTCGAC
G G W L Q N L N S L S L D F S Y A A D -
ACGGTATCTCATTTGACTACACAAAGTATCACCTCGCGCTCTTGGTAAGACTCGTGTGATGAAC + 4080
TGCCATAGAGTAACCTGATGTGTTTCATAGTGGAGCGCGAGAACCATTCCTGAGCAGCTACTTG
G I S L T T Q V S P R A L G K T R D E Q -
AAGTGTATACTTGGTAACAATCCTTGATGGTTACTTCGAAAACGGTGGACAAACAGCTTA + 4081
TTCAACTANTGAACCATGTTAGGAACCTACCAATGAAGCTTTTGCCACCTGTTGTGCAAT
V D N L V T I L D G Y F E N G G Q H V N -
ACTTGAACGTTATGGACTTGAACGATGTTTACGAAAAAATCATGTCAGGCGAAGACGTTA + 4141
TGAACCTGCAATACCTGAACCTTGCTACAAATGCTTTTATAGTACAGTCGCGCTTCTGCAAT
L N V M D L N D V Y E K I M S G E D V I -
TCGTACGTATCTCTGGATACCTGTATAACACATAAATACCTCCTCCTCAGAACAAAACTG + 4201
AGCATGCATAGAGACCTATGACACATTTGTGATTTATGGAGTGAAGTCTGTTTGTGAC
V R I S G Y C V N T K Y L T P E Q K T E -
AATTGACACAACGCTGCTTCCAGGAAGTTCCTTCAATGGATGACGCTTGGATGCAATTGA + 4261
TTAACTGTGTTGCACAGAAGGTGCTTCAAGAAAGTTACCTACTGCGGAACCTACGTAAC
L T Q R V F H E V L S M D D A L D A L S -

261/487

Figure 101J

4321 GCTAATCAAGTCTTGAATAATAAAGGGCTCTTTGTCAACTGTAGTGGTGAAGAAA 4380
-----+-----+-----+-----+-----+
CGATTAGTTCAGAAGCTTATTTATTTTCCCGAGAAACAGATTGACATCACCCAACTCTCTTT
c *
4381 AGCTAAGCTCGAGAAAGGACAAATTTGTCCCTTCTTTTGTATGTTTCAGAGCGATGAAA 4440
-----+-----+-----+-----+-----+
TCGATTCGAGCTCTTTCCTGTTTAAACAGGAAAGAAAAAACTACAAGTCTCGCTACTTT
e * A R S L V F K T R E K K I N L A I F -orf1_670 homologue of sp0460, transposase
4441 ATCCGTTTTTTGAAGTTTCAAAAGTTCCGAAACCAAGGCAATTCGCTTGATGCTTTTG 4500
-----+-----+-----+-----+-----+
TAGGCAAAAAAAGTCAAAAGTTTCAAGGCTTTTGGTTTCCGTTACCGGAACACAGAAAAC
e I R K K F N E F N R F G F A N R K I D K -
4501 ATGAGTTTGTAGTGGCTCAAGTTTAGCGTTAGATAAAGGCAATCAATGGCGTTAGTG 4560
-----+-----+-----+-----+-----+
TACTCAACAAATCACCGGAGTTCAATCGCAATCTTATTCCTGTTAAGTTACCGCAATCAC
e I L K N T A E L K A N S Y P L E I A N T -
4561 ATGTAGTTTTTATAGCAATAAATGTGCTCAAAGTGGTTTAAAGGTGCGGTTGAGATGA 4620
-----+-----+-----+-----+-----+
TACATCAAAATATCGTTTATTTACACGAGTTTCACCAAAATTTCCACGCCAACTCTACT
e I Y N K Y C I F T S L T T K F T R N L H -
4621 GGTAACGTGCTTGAATTAAGCCCAAACTGGTCAGTATTCCTCTTGTAGATGAAAT 4680
-----+-----+-----+-----+-----+
CCATTGCACAGAACTTAATTCGGGGTTTGGACCAATCATAAGAGAGACATCTACTTTA
e P L T D Q I L G W F Q D T N K E Q L H F -
4681 AGGAGTAGTGTATACAGTCAATAGTAATCTTTAAGTTCAGGTACTAGAGTAAAGATTTTC 4740
-----+-----+-----+-----+-----+
TCCTCATCAACTATGTCAGTATCATTAGAAATCAAGTCCATGATCTCATTTCTAAAG

262/487

Figure 101K

e L L L Q Y L D Y Y D K L E P V L T F I K -
4741 TTGAGACACTCCCTAGGAGTTAAGGTCCTCTGTGAAAGTTCTAGCATAGAAAGGCTTAAGA + 4800
AAGTCGTGAGGGATCCCTCAATTCAGAGACATTCAGATCGTATCTTTCGGAATTC
K L C E R P T L T E R F T R A Y F P K L -
4801 GAGAGTTCCGACTATCTTTAGGATAAATTTCCAGTAATATTTAAGAGCCTGTGATTCC + 4860
CTCTCAAAGGCTGATAGAAAATCCCTATTAAAGGTCATTATAAATTCGAGACATAAGG
S L K R S D K L I F K W Y Y K L A R Y E -
4861 AGAGATTATCATCAAAATTCCTTCATGATGTTGATTCTAGTCTGATTAAGACCCCTGCTC + 4920
TCTCTAAATAGTAGTTAAAGGAAGTACTACAACCTAAGATCAGACATAATTCGCGGACGAG
L S K D D F Q K M I N I R T Q N L A R S -
4921 ATGTTGTTGGACAATGTGGAACGATCGAGAACAAATTTAGCATGGGAAATAATTTCTTA + 4980
TACACAACCTGTTACACCTTTGCTAGCTCTGTTGTTAAATCGTAACCTTTTATTAAGAAAT
M H Q V I H F R D L V I K A N P F L K K -
4981 ATGAGAGGGATATAACTTCCAGACATATCAACAGTACGACTTAACTTTTCTTCTAGCT + 5040
TACTCTCCCTATATTGAAGGCTCTGTATAGTTGTCACCTGCTGTAATGAAAAAAGATCGA
I L P I Y S G S M D V T V V K V K K R A -
5041 TCTTTCGAGTACTTGAGAAATGATTTCGGATGGTTGTTTGGACGCTCTGTATCAAGAATG + 5100
AGAAAGCTCATGAACCTTCTTACTAAAGCCCTACCAACAACTGCAGACATAGTTCTTAC
E K S Y K F F H N R I T T Q R R N D L I -
5101 GTCATGATTTTCTTAGTGTGAAATCCTGAGCAATGAAGCAATTTCCCTTCTGCTAG + 5160

263/487

CAGTACTAATAAGAAATCACAACCTTTAGGACTCGTTACTTTCGGTTAAGGGGAAGACCATC

Y
Y
K
G
K
L
A
F
I
A
Q
D
F
N
T
K
K
I
M
F
e

5161 GAGAAATTCATCCCGAGGAGGATTTTCAGGCCAAAGTGGTGTAATCTCTTGGAAATGAAAT
-----+-----+-----+-----+-----+-----+-----+
CTCTTAAGTAGGGTCCCTCTCTTAAAGTCCGGTTTCAACACATTTAGGAGAACCTTTTACTTTTA
5220

e S F E D W S L I E P L T T Y D E Q F H F -

5221 TGCTTGAGCTTACGATACACGGTAGAGGTAGAGGTAGAGATGGCTAATTTAGAAAGCGATA
5280 ACGAACTCGAATGCTATCTGTGCCATCTCCATCTCCATCTCTACCGATTAAATCTTCGTAT

e e Q K L K R Y V T S F S I A L K S A I -

5281 TGTGTAAGAGCCCTCTCTGTGTTGAGTAGGAGTTGGCAATTTTTGTCTCACCATTTCCGAG
-----+-----+-----+-----+-----+-----+-----+-----+
ACACATTTCTCGGAGAGACAACTGATCTCTCAACCGGTTAAAGACACAGATGGTGAAGAGCTC
5340

THE L A E R N L L Q A I K Q R V M E S -

5341
ATTTGGCAATTTTCTGAACGAGAGTTGTTTCAGCTACAGTCACTTTCGCACAGGACTTG
-----+-----+-----+-----+-----+-----+-----+-----+
TAAACCGTTAAAAAGACTTGCTCTCAACAAAGTCGATGTCACTGAAAGGCGTGTCTCTGAAC

Q C N K Q V L T E A V T V K R C S K

5401 CATGAAATCGTCTCTTTTCAATGAATAGGCTAGGGAACCAATCTCGATAAAA
-----+-----+-----+-----+-----+-----+
GTAACTTTAGCAGAGAAAAAGTTTACTTACTCCGATCCCTTGGTGGTAGAGCTATTTT

Q
C
Q
E
R
R
K
K
L
H
H
L
S
A
F
G
G
H
H
E
L
L

5461
GGGATTTTAGAAGGCTTTTGGAAACGCTATTTTGATTTTCTTTACAGTGTTACAT
-----+-----+-----+-----+-----+-----+
CCCTAAATCTCCGAAAACCTTCAGCATAACTAAACAAAAGGAATGTCAACAANGTA
-----+-----+-----+-----+-----+-----+ 5520

a e p i k s p k q f d y k i o k g k c h k c -

Figure 101M

TTAGTGGTGATATCAAGTGTAGCGAAGACTTCGATATGGGTATCGTCTGAATGGCT 5521
-----+-----+-----+-----+-----+-----+ 5580
AATCCACCCACTATTAGTTCACATCGCTTCTGAAGCTATACCCATAGCAGGACTTACCGA
K P P H Y D L T A F V E I H T D H Q I A -
TTATTTAAGGTGATGTTTTTGTCTTTTATTCGATGAGTAATGTGGTATGATGTGT 5581
-----+-----+-----+-----+-----+-----+ 5640
AATAAATCCCACTACAAAAACAGAAAAATAAGGCTACTCATTACACCACTAATACTACACA
K N L T I N K D K I G I L L T T H N I H -
TCCATAAGATACTTTTCTAATGAGTTGTTTAGGGCTTTTTCATTATATAAGTCTTATGGGACT 5641
-----+-----+-----+-----+-----+-----+ 5700
AGGTATTCATGAAAGATTACTCAACAAATCCCGGAAAGTAATATTCAGAAATACCCCTGA
AGGTATTCATGAAAGATTACTCAACAAATCCCGGAAAGTAATATTCAGAAATACCCCTGA
E M
TTTTTGATACTCAAAAAGCCCTTAATACTCCACAGTGGGATTACCCACTACAGAAATTA 5701
-----+-----+-----+-----+-----+-----+ 5760
AAAACTATGAGTTTTTCGGGATATTAGAGGTGCACCCCAAATGGGTCATGCTTTAAT
TAGAGCCAGAAAAACACTTTTGTCTACTAGCAGAAACTAGAGAGCAGAGTGTTTTCT 5761
-----+-----+-----+-----+-----+-----+ 5820
ATCTCGGTCCTTTTGTGAAAAACAAGTGATCGTCTTTGATCTCTCGTCTTCACAAAAAGA
GTTCAGATTTACCCAAAACTGGGAAATATGGGGATAAGAATAGAGATGGCTTAGGAAGCC 5821
-----+-----+-----+-----+-----+-----+ 5880
CAAGTCTAAATGGGTTTGTACCCCTTATACCCCTATCTCTTATCTCTACCGAATCCTTCGG
CCTTTTGTGTGTAGACAGTACGATGAACCTTATAACAAATAGTGAGCCCTTTTACGAATC 5881
-----+-----+-----+-----+-----+-----+ 5940
GGAAAAACACAGATCTGTCTACTTGTGAATATTGTTTATCTACTCGGAAAAAATCGTTAG
* L L Y H A K K A I -orf2_670 homologue of sp0461, transcriptional
regulator
ATTGCCACCCGTTTGTCAAAAGCCCTCTTTTTCGGATATCTACAATGTCTGATAGATGAGA

```

5941 -----+-----+-----+-----+-----+-----+-----+-----+ 6000
TAACGCTGGCAAAACAGTTTTCGGAGAAAAGCCTATAGATTGTAAACAGACATCTACTCT
e      M A V R K D F A E K R I D V I T Q Y I L -

CGTGTTGGCTAACATGAATACTAAGCCAATCGTCAA AAAAGTGATGTTCCCTTTGGGA
6001 -----+-----+-----+-----+-----+-----+-----+-----+ 6060
GGCAACAACCGATTGTACGTTTAGATTCCGTTAGCAGTTTTCACACAAAAGGGAACCGCT
R Q Q S V H L D L A I T L F H K G K P -

TACTGCTTTTTTAACGTAAGGCAGGTATTCTTTTCGTTGTATAATAATCAATGGCTCTGTC
6061 -----+-----+-----+-----+-----+-----+-----+-----+ 6120
ATGACGAAAAATGCAATTCGTCATAGAAGAAACAATATTATTAGTTACCAGAGACAG
Y Q K K V Y P L Y E K T T I I I L P E T -

AAATGCTCCTCTGAAGGAGGAGGACTAATTAGAATATTCTATCTCTGAACAGAGGCAACT
6121 -----+-----+-----+-----+-----+-----+-----+-----+ 6180
TTTACGAGGAGACTTCCTCCTCCTGATTATCTTATAACATAGGACATTTGTCCTCGTTGA
L H E E S P P S I L I N Y G T V S A V -

TTTTCAGTAAAAATTCGGTAAAAATATGGACITTTATTAACTTTACATCTGCTGATTATTT
6181 -----+-----+-----+-----+-----+-----+-----+-----+ 6240
AACAGTCATTTTAAGGCATTTTATTACCTGAAATAATTCAAATGTAGACGAACATAATAAA
K D T F N R L I I S K I L N V D A Q N N -

AAAATGATAAAAATCGGATAGCAGTAGTGAGGAAAAGATGGTTTCTGTCAAGTAGAGT
6241 -----+-----+-----+-----+-----+-----+-----+-----+ 6300
TTTTACTATTTTAGCCCATGCTGCCATCACCTCCTTTTCTACCAAAAGACAGTTCATCTCA
L I I F I P I A P L S S F I T E T L Y L -

GAGAAAAGGTACAGCCGATGCTGGTCGATAACTCCTTCAATCTTCTGCTCAGTCATCCAC
6301 -----+-----+-----+-----+-----+-----+-----+-----+ 6360
CTCTTTTCCATGTCGGCTACGACCGACTATTGAGGAAGTTAGAAGCAGGAGTCAGTAGGTG
S F L Y L R H Q D I V G E I K Q E T M W -

```

PCT/US05/27239

266/487

Figure 1010

```
6361 TCCTGAACAATTGCTTTTCGAAATATGATACAGTGGCTTGTGCTTTCATCCCATTAATG 6420
-----+-----+-----+-----+-----+-----+-----+
AGAACTTGTTAACGAAAGCTTTATACATGTCACCGACAGCGAAGTTAGGGTATTACA
e E Q V I A K S I H Y L P K D S E I G Y H -
6421 TCGTAATAATTATAATAGGGAAGTATGTTGTAACCAAAACAAACGTTCTTGTTAAG 6480
-----+-----+-----+-----+-----+-----+-----+
AGCAATTAATAATATATCCCTTGATCTAAACATTTGGTTTGTGTTTGCAGAACAAATTC
e E Y Y N Y Y P V L N Q L G F L F T R T L -
6481 AAATCAGTGTCTTAAAGAAAGAGAAATTCGAATGTCATTTCTTAAGATATCTTG 6540
-----+-----+-----+-----+-----+-----+-----+
TTTCAGTCACGACAATTTTCTCTTAAGCTTTACAGTAAAGGATTCCTATAAGAAC
e F T L A T L F S L S N S I D N G L I N K -
6541 AACTGGATAGTAGATGCTTTCCTCTTGTATGCTGAAGAATCAGTGAATAGTATGATC 6600
-----+-----+-----+-----+-----+-----+-----+
TTGAACCTATCATCTACGAAAGGAGAACATACGACTTCTTAGTCACTTATCATCTACTCAG
e F K S L L H K G R T H Q L I L Q I T H T -
6601 TTTTTCCTTGATTCATTTTGTCTTGGAAAACGAAGAATTAGCAGAACATAAACCAA 6660
-----+-----+-----+-----+-----+-----+-----+
AAAAAAGAACTAAGGTAAACAGGAACCTTTGCTTCTTAATCGTCTTGTATTGTTT
e K K E Q N W K D K S F S S N A S C Y V L -
6661 AAGATATAATCCAGTTCCTCTGAGTAAAGTCATGTTGGCATGTTGGCTCTAAGTAAGTT 6720
-----+-----+-----+-----+-----+-----+-----+
TTCTATATTAGGTCAAGAAGGACTCATTTTCAGTACACCGTACACCGAGATTCATCAA
e F I Y D L E E Q T F T M N A H P E L Y T -
6721 TGGCAATGTTCCATCAAAATCGGATACATAAAGAGTTTAAATTTTCAAACTCTTTG 6780
-----+-----+-----+-----+-----+-----+-----+
ACCGTTACAAGTAGTTTACCTATGTTATTTCTCCAAAAAATTAAGTTTGAGAAAC
```

Figure 101P

[illegible]

268/487

Figure 101Q

7201
e G Q L L F F R L I K L I T S Q N Y L Q H -
GGAACATCATCTTCTTGGCCCTCATAAAATTTTATCAACTAACCAATATTTGACTACC
AAGTAATAATTCTGTTGATGAGAAATGGTTCGATTAAATGAACCTGTTGCGTATCTAAA
TTCTATTATTAAGCAAACTACTCTTACCACAGCTAATTAACCTTGAAACAACGCATAGATTT
F Y Y N T Q H S H E I L Q V Q Q T D L -
TTAAATGTCACACTCTTCCGAAATGTTTCTGTGAATTCCTGCAAAATGCTTAGGAGACTT
AATTTACAGTTGAGAAGGAGCTTACAAGAACATTAAGGACGTTTACGAATCCTCTGAA
N F T L E E E F T E Q L E Q L I S L S -
TTAGATTGTAATGAAGTTAAAGTAGACAGTTCATCTAGTCAATAGACCGAATATCCAAT
AATCTAACATTACTTCAATTTTCATCTGTCAAGTAGATCAAGTTATCTGGCTTATAGGTTA
K S Q L S T L T S L E D L E I S R I D L -
AATATATTATAAAGGTAATTTTATCTGTGAATCTTTTTCATCTGATTTGTTTAGCATA
TTATATATAAATTTTACCATTAAATATAGACATTAAGAAAAGTTACATAAACAAATCGTAT
L I N L I T I K D T I R K E I Y K N L M -
GTACCGAATCTTAGTTGCATATAGATAATTTTAATTAATTAATACAAAAGAACTAAT
CAATGGCTTAGAATCAACGTTATCTATTAAAAATTAATAATATTATGTTTTCTTTGATTA
TGCTCTGTCAAAAAGGTTGTGGAATTTCCGACTTTTATTGATAAAACAGCATGTAATAAAA
ACAGAACAGTTTTCCTCCACACCTTAAAGGCTGAAATACTATTTTGTGCTACATATTTT
GGCATTTTAAAGATAGTAATGAGTATTTGGTGGAGTTTATGCTTATTTTTTTTATTAGA
7260
7320
7380
7440
7500
7560
7620

Figure 101R

CCGTAAATTTCTATCATTTACTCATTAACCCACCTCAAAATACCGAATAAAAAAATAATCT
7621
AAATATTTTTTTTATCAAAATATTGTCGTTCTATATAAAAAATATGATGATAAAAAATATCTATT
7680
TTTATAAAAAAATAGTTTATATAACAGCAAGATATTTTTTTTATACACTATTTTATATAGATAA
GTGATGGAAGTGTGTTTAAATTTATACTAGGATAGTAAATAGTAATACTATACATATCTAT
7740
CACTACCTTCAACAAATTAATATGATCCTATCAATTTATCATTTATGATATGATATGATA
ATTGTATACAAAGTGTGTCATGTCAGGTTGAGAAAGATAGCTATAACGCACCTTTTATACGC
7800
TAACATATGTTACACACATAACGGTCCAACTCTTCTATCGATATTGCGTGAATAATGCG
TTTTGCTACGTTTGTAGTGAACGGATTAACTCAGTGAGATAAATTTTATCAGAACATAA
7860
AAAACGATGCAAAACAATCACTTGCCCTAAATGAGTCACTCTATTATAAAATAGTCTTTGTATT
GTAATCCGTTTCTTCGTTGATACAGATTGAAAGTACCTATGATCATAGAGGATTAACT
7920
CATTAGGCAAGAAGCACATATGTCATACTTTTCATGGATACCTAGTATCTCTCCTAATTGA
TGTTCTATGAATAATGCTTAACAGGAGACACACATGAAAAAGTAAGAAGATATTCA
7980
ACAAGATACTTATTACGAATTGTCCTCTGTTGTTGTTCTTTTTCATTTCTTCTATATAAGT
GAAGGCAGTTGCAGGACTGTGCTGATATCTCAGTTGACAGCTTTTCTTCGATAGTTGC
8040
CTTCGGTCAACGTCCTGACAGCACATATAGAGTCACTGTGCGAAAAAGAGCTATCAACG
b M L N R E T H M K K V R K I F Q -orf3_670 homologue of sp0462, LPXTG
K A V A G L C C I S Q L T A F S S I V A -

```

8041      TTTAGCAGAAACGCCCTGAAACCCAGATCCAGCGATAGGAAAGTAGTGATTTAAGGAGACAGG
      +-----+-----+-----+-----+-----+-----+
      AATTCGTCTTTGCGGACATTGGTCAGGTCCGCTATCCCTTTTCATCACTAATATTCCTCTGTC
      L A E T P E T S P A I G K V V I K E T G -
      CGAAGGAGGAGCGCTTCTAGGAGATGCGCTGTTGAGTTGAGTTGAAAAACAATACGGATGGCAC
      +-----+-----+-----+-----+-----+-----+
      GCTTCCTCCTCCGGAAGATCCTCTACGGCAGAACTCAACTTTTGTGTTATGCGCTACCGTGC
      E G G A L L G D A V F E L K N N T D G T -
      AACTGTTTCGCAAAAGGACAGAGGCGGCAACACAGGAGAGCGATATTTTCAACATAAAACC
      +-----+-----+-----+-----+-----+-----+
      TTGACAAAAGCGTTTCCTGCTCTCCGCGTTGTCTCTCTTCGCTATAAAAGTTTGTATTTTGG
      T V S Q R T E A Q T G E A I F S N I K P -
      TGGGACATACACCTTGACAGAAGCCCAACCTCCAGTTGGTTATATAAACCTCTACTAAACA
      +-----+-----+-----+-----+-----+-----+
      ACCCTGTATGTGGAACCTGCTCTCGGGTTGGAGGTCAACCAATATTTGGGGAGATGATTTGT
      G T Y T L T E A Q P P V G Y K P S T K Q -
      ATGGACTGTGTAAGTTGAGAAGAAATGGTCGGACGACTGTCACCAAGTGAACAGGTAGAAAA
      +-----+-----+-----+-----+-----+-----+
      TACCTGCAACACTTCAACCTCTCTTACACAGCCTGCTGCAGAGTTCCACTTGTCCTCATCTTTT
      W T V E V E K N G R T T V Q G E Q V E N -
      TCGAGAAGAGGCTCTATCTGACCACGATCCACAACACAGGGAAGTATCCAGATGTTCAAC
      +-----+-----+-----+-----+-----+-----+
      AGCTCTTCTCCGAGATAGACTGGTCAATAGGTGTTGTTGTCCTGTAATAGGTCTACAAAGTTTG
      R E E A L S D Q Y P Q T G T Y P D V Q T -
      ACCTTATCAGATATTATAGGTAGATGGTTCCGAAAAAAAACGGACAGCACAGCGGTTGAA
      +-----+-----+-----+-----+-----+-----+
      TGGAAATAGCTTAATAATTCACATCTACCAAGCCTTTTTCCTGCTGCTGTTCCGCAACTT
      +-----+-----+-----+-----+-----+-----+
8401

```


PCT/US2005/027239

271/487

Figure 101T

b P Y Q I I K V D G S E K N G Q H K A L N -
TCCGAATCCATATGAACGTGTGATTCCAGAGGTACACTTTCAAAGAGAATTTATCAAGT 8520
8461 +-----+
AGGCTTAGGTATACCTGCACACTAAGGTCTTCCATGTGAAAGTTTCTCTTAATAATAGTTCA
P N P Y E R V I P E G T L S K R I Y Q V -
GAATAATTGGATGATAACCAATATGGAATCGAGTTGACGGTTAGTGTAATAAACACGCGGT 8580
8521 +-----+
CTTATTAAACCTACTATTGGTTATACCTTAGCTCAACTGCCAATCACCATTTTGGTGTCCA
N N L D D N Q Y G I E L T V S G K T T V -
TGAAACGAAAGAAGCCTCTACTCCGCTAGATGTGTGTTATTCTATTAGATAACTCCAATAG 8640
8581 +-----+
ACTTTGCTTTCTTCGGAGATGAGCGGATCTACAACAATAAGATAATCTATTGAGGTTATC
E T K E A S T P L D V V I L L D N S N S -
TATGAGTAATATTCGACATAATCATGCCCCATCGAGCGGAAAAAGCGGAGAAAGCGACACG 8700
8641 +-----+
ATACTCATATATAAGCTGTATTAGTAGGGGTAGCTCGCCTTTTTCGCCCTCTTCCGCTGTGC
M S N I R H N H A H R A E K A G E A T R -
AGCCCTTGTAGATAAGATTACCTCCAATCCAGATAATCGAGTAGCAGTGTGACTTATGG 8760
8701 +-----+
TCGGGAACATCTATTCTAATGGAGGTTAGTCTATTAGTCTATCGTGAACACTGAATACC
A L V D K I T S N P D N R V A L V T Y G -
CTCAACTATCTTTGACGGTTCAGAAGCTACTGTGGAAAAAGGGGTAGCAGATGCGAACGG 8820
8761 +-----+
GAGTGTAGAGAAACCTGCCAAGTCTTCGATGACACCTTTTCCCCCATCGTCTACGGCTGCC
S T I F D G S E A T V E K G V A D A N G -
AAAAATATTGAATGACTCAGCTTTATGGACGTTTCGATCGTAGACGTTTACAGCTAAAC 8880
8821 +-----+

Figure 101b

TTTTTATAACTTACTGAGTCGAAATACCTGCAAGCTAGCATGCTGCAAAATGCTGATTTTG
K I L N D S A L W T F D R T T F T A K T -
TTATAATTATAGCTTTTAAATCTCACATCAGATCCTACTGATATTCAAACTATTAAGGA 8940
8881 +-----+
AATATTAAATATCGAAATTTAGAGTGTAGTCTAGGATGACTATAAGTTTCATATAATTCCT
Y N Y S F L N L T S D P T D I Q T I K D -
TAGGATTCCATCAGATGCAGAGGAATTGAACAAAGACAAAATGATGATATCAATTCGGGCG 9000
8941 +-----+
ATCCTAAGGTAGTCTACGTCCTCTTAAGTTGTTTCTGTTTAACTACATAGTTAAGCCGCG
R I P S D A E E L N K D K L M Y Q F G A -
GACTTTTACCCAGAAAGGCTTTGATGACCGCTGATGATATCTTGACAAAGCAGGCAAGACC 9060
9001 +-----+
CTGAAATGGGTCTTCCGAAACTACTGCGGACTACTATAGAACTGTTTCGTCCTGTTCTGG
T F T Q K A L M T A D D I L T K Q A R P -
AAACAGTAAAGGTTATTTCCACATTACAGATGGTGTTCGACTATGTCATATCCAAAT 9120
9061 +-----+
TTTGTCATTTTCCAAATAAAGGTGTAATGCTCTACCACAAAGGCTGATACAGTATAGGTTA
N S K K V I F H I T D G V P T M S Y P I -
TAATTTTAAATATACAGGAACGACGCAATCGTACAGAACTCAGCTGAATATTTTAAAGC 9180
9121 +-----+
ATTAAATTTATATATGTCCTGCTGCTGCTAGCATGCTGTGAGTCTGAGTCTATATAAATTCG
N F K Y T G T T Q S Y R T Q L N N F K A -
AAAACTCCAAATAGTAGCGGGATATTAAGTGGAGGACTTTGTTACATGGTCAGCAGATGG 9240
9181 +-----+
TTTTTGAGGTTTATCATCGGCCCTATATGACCTCCCTGAAACAAATGTACCACTCGTCTACCC
K T P N S S G I L L E D F V T W S A D G -

PCT/US05/27239

273/487

Figure 101V

9241 TGAACATAAGATTGTTCTGGTGGAGATGGTGAAGTTATCAGATGTTTACGAAGAAACCTGT
-----+-----+-----+-----+-----+-----+ 9300
ACTTGTATTCTTAACAAGCACCTCTACCACTTTCATAATAGTCTACAAATGCTTCTTTGGACA
b E H K I V R G D G E S Y Q M F T K K P V -
9301 AACAGACCAATACGGAGTTTCATCAAAATACCTTCAATCACCTCCATGGAGCAGAGAGCTAA
-----+-----+-----+-----+-----+-----+ 9360
TTGTCTGGTTATGCCCTCAAGTAGTTTATGAAGTTAGTGGAGGTACCTCGTCTCTCGATT
b T D Q Y G V H Q I L S I T S M E Q R A K -
9361 ATTAGTTTCAGCGGGATATAGTTCTATGGAAGTACTGACTTGTATTATTATTTGGCGTGATAG
-----+-----+-----+-----+-----+-----+ 9420
TAATCAAAGTCGCCCTATATCCAAAGATACCTTGACTGAACATAAAATAACCGCACATATC
b L V S A G Y R F Y G T D L Y L Y W R D S -
9421 TATTCTAGCCATATCCATTAACTCTAGTACCGATGGATTGATTACCAACCATGGTGACCCCTAC
-----+-----+-----+-----+-----+-----+ 9480
ATAAGATCGGATAGGTAAATTGAGATCATGGCTAACCTAATGGTTGGTACCACACTGGGATG
b I L A Y P F N S S T D W I T N H G D P T -
9481 GACTTGGTATTATAACGGAAATATGGCTCAGGATGGCTATGATGTCTTCACCTGTTGGGT
-----+-----+-----+-----+-----+-----+ 9540
CTGAACCATATAATTGCCCTTTATACGGAGTCCCTACCGATACACAGAGTGACAAACCCCA
b T W Y Y N G N M A Q D G Y D V F T V G V -
9541 TGGTGTAAACGGGATCTGGTACGGATGAAGCAACGGCTACTAGATTATTATGCAGAGCAT
-----+-----+-----+-----+-----+-----+ 9600
ACCACATTTGCCCTTAGGACCATGCTACTTCTGTTGCCGATGATCTAAATACGTCCTCGTA
b G V N G D P G T D E A T A T R F M Q S I -
9601 CTCTAGTTCTCCTGACAACTACACTAACGTAGAGATCCATCTCAGATTTTACAGAATT
-----+-----+-----+-----+-----+-----+ 9660
GAGATCAAGAGGACTGTTGATGTGATTGCATCGTCTAGGTAGAGTCTAAATGTTCTTAA

Figure 101W

b S S S P D N Y T N V A D P S Q I L Q E L -
9661 GAATCGCTACTTCTATATCTGTCATGAGAGAAATCTATCGAAATGGTACGATTAC 9720
CTTAGCGATGAAGATATGATAGCAGTTACTCTCTCTTTAGATAGCTTTACCATGCTAATG
N R Y F Y T I V N E K K S I E N G T I T -
9721 AGACCGGATGGGGAACCTAATTGATTTCCAAATGGGAGAGATGGAGGTTTGTATCCAGC 9780
TCTGGGCTACCCACTTGATTAACCTAAAGGTTAACCCCTCGTCTACCTCCCAACATAGGTCG
D P M G E L I D F Q L G A D G R F D P A -
9781 GGATACACTTTAACTGCAACGATGGTAGTTCGTTGGTCAATAATCTCCCTACTGGGGG 9840
CCTAATGTGAATTGACGTTTGTCTACCATCAAGCAACCACTTATTACAGGGATACCCCC
D Y T L T A N D G S S L V N N V P T G G -
9841 ACCACAAAATGATGGTGGCTTGTCTAAATAAATGCAAAAGTCTTCTATGATACGACTGAGAA 9900
TGGTGTTTACTACCAACCGAAGCATTTTCTACGTTTTCACAAGATAGTACTGCTACTCTT
P Q N D G G L L K N A K V F Y D T T E K -
9901 AAGATCGTGTACAGGTTTGTACCTTGGAAACGGGTGAAPAAAGTTACATTGACTTATAA 9960
TTCCTAAGCACATTTGTCCTCAACATGGAACCTTGGCCACCTTTTCAATGTAACATATTT
R I R V T G L Y L G T G E K V T L T Y N -
9961 TGTTCGTTGAATGACCAATTTGTAAGCAATAAATCTATGACACGAAATGGTCGAACAAC 10020
ACAAGCGAACTTACTGGTTAAACATTCGTTATTAAAGATAGTGTGCTTACCAGCTGTGTG
V R L N D Q F V S N K F Y D T N G R T T -
10021 CCTACACCCCTAAGGAAGTAGAAAAGAACACAGTGGCGGACTTCCCGATTCCCTAAGATTGCG 10080

Figure 101X

GGATGTGGGATTCCTTCATCTTTCTGTGTCAACGGCTGAAGGGCTAAGGATTCCTAAGC
L H P K E V E K N T V R D F P I P K I R -
10081
TGATGTACGAAAGTATCCAGAAATCACAATTCCTCAAAAGAGAAAAACCTGGTGAAATGA
ACTACATGCTTTCATAGGCTTTAGTCTTTAGGTAAAGTTTCTCTTTTGTGAACACACTTTAACT
D V R K Y P E I T I P K E K K L G E I E -
10141
GTTTATTAGATCAATAAGAAATGATAAAAAACCACTGAGAGATGCCGCTTTAGTCTTCA
CAAATAATCTAGTTATCTTACTATTTTGGTACTCTCTACGCCAGAAATCAGAACT
F I K I N K N D K K P L R D A V F S L Q -
10201
AAAACAACATCCGGATTATCCAGATATTTATGGAGCTATTGATCAAAAATGGCACCTTATCA
TTTTGTTAGGCCCTAATAGTCTATAAATACCCTGATACTAGTTTACCGTGATAGT
K Q H P D Y P D I Y G A I D Q N G T Y Q -
10261
AAATGTGACAAACAGGTGAAGATGGTAAGTTGACCTTTAAATCTGTCTAGATGGGAATA
TTTACACTCTGTGCCACTTCTACCAATCAACTGGAAATTTTATAGACAGCTACCCCTTAT
N V R T G E D G K L T F K N L S D G K Y -
10321
TCGATTATTGAAAAATCTGAACCACTGGTTATAAACCCGTCAAAATAGCCCTATCGT
AGCTAATAAACCTTTTAAGACTTGTGACCAATATTTGGGCAAGTTTATTCGGATAGCA
R L F E N S E P A G Y K P V Q N K P I V -
10381
TGCCCTCCAAATAGTAATGGAGAGTCAGAGATGTGACTTCAATCGTTCACAAAGATAT
ACGGAAGGTTTATCATTTACCTCTTACCTCTCTACACTGAAGTTAGCAAGGTGTTCTATA
A F Q I V N G E V R D V T S I V P Q D I -

Figure 101Y

10441 ACCAGCGGTTACGAGTTTACGAATGATAAGCACTATATATCACAATGAGCCAAATTCCTCC + 10500
-----+-----+-----+-----+-----+-----+-----+-----+
TGTCGCCCAATGCTCAAAATGCTTACTATTCGTGATATAGTGTCTTACTCCGGTTAAGGAGG
b P A G Y E F T N D K H Y I T N E P I P P -
-----+-----+-----+-----+-----+-----+-----+-----+
10501 AAAAAAGAAATATCCTCGAACTGGTGGTATCGGAATGTGCCATTCTATCTCATAGGTTG + 10560
-----+-----+-----+-----+-----+-----+-----+-----+
TTTTTCTCTTATAGGAGCTTGACCACCATAGCCCTTACAACGGTAAAGATAGACTATATCCAAC
K R E Y P R T G G I G M L P F Y L I G C -
b -----+-----+-----+-----+-----+-----+-----+-----+
10561 CATGATGATGGGAGGAGTTCTATATATATACACACGGAACATCCGTAAAGTGTAGCAATGAG + 10620
-----+-----+-----+-----+-----+-----+-----+-----+
GTAATACTACCCCTCCTCAAGATAATATGTGTGCCCTTTGTAGGCATTTTCACATCGTTACTC
M M M G G V L L Y T R K H P *
b -----+-----+-----+-----+-----+-----+-----+-----+
10621 AAATGATAATATCGATACTCTGAGCGGATATCTTTTAAAGTAGTACACTCAAGAGAGATTT + 10680
-----+-----+-----+-----+-----+-----+-----+-----+
TTTACTATTATAGCTATGAGACTCGCTATGAAATTTCTTCATCGTGAGTTCTCTCTCTAAA
AAGTTTACTTGGTGAACACAGTTTCTTCGCCAAGTAAACACCATTGAAAGGGGAGATG + 10740
-----+-----+-----+-----+-----+-----+-----+-----+
TTCAAATGAACCACTTTTGTCAAAGAGCGGTTCAATTTGGTGGTAACTTTCCCTCTCTAC
TTTTTCGAAACTTGCACAGAAAAAGGATTTATTTCTCATGTGTAAATTCATTACATTGC + 10800
-----+-----+-----+-----+-----+-----+-----+-----+
10741 AAAAGCTTTTGAACGTGCTTTTTTCCCTAATAATAACAGTACACATTAAAGTAATGTAACG
-----+-----+-----+-----+-----+-----+-----+-----+
TCACAGTTGATTTTAAAGAGATATGAATAAGGAGAAATCATGAATCAATCAACAAATTTT + 10860
-----+-----+-----+-----+-----+-----+-----+-----+
10801 AGTGCAACTAAATTCCTCTATACTTATTCCTCTTTTACTACTTTTACTAGTTGTTTAAAA
-----+-----+-----+-----+-----+-----+-----+-----+
M K S I N K F L - orf4_670, homologue of sp0463, LPXTG
c TAACAATGCTTGCTGCCCTTATATCTGACACGAGTAGCCTGTTTTCAGCTGCAACAGTTT

[illegible]

PCT/US05/27239

278/487

Figure 101AA

```
11281 TTGGTCCTAATGGGCAAGTATTAAACAGTTCAAAAGCGGTACCTGCTCTTGTAACTCTTC + 11340
-----+-----+-----+-----+-----+-----+-----+
AACCAGGATTACCCGTTCAATTAATTGTCCAAGTTTTCGGCATGGACGAGAACATTGAGAAG
      G P N G Q V L T G S K A V P A L V T L P -
11341 CACTTGTTAACAAATAATGGTACAGTAATTGATGCACATGTTTCCCTAAATAATTCATATA + 11400
-----+-----+-----+-----+-----+-----+-----+
GTGAACAATTGTTATTACCATGTCTATTAACTACGTGACAAAAGGGATTTTAAAGTATAT
      L V N N N G T V I D A H V F P K N S Y N -
11401 ATAAACAGTTGTAGATAAAAAGAATTGCTGATACCTTTGAATTATAACGATCAAAATGGTC + 11460
-----+-----+-----+-----+-----+-----+-----+
TATTGGTCAACATCTATTTTCTTAACGACTATGAACCTTAATATGCTAGTTTACCAG
      K P V V D K R I A D T L N Y N D Q N G L -
11461 TGTCTATCGGTACTAAAATCCCATATGTTGTTAATACAAACAATTCCTCAAGTAATGCAACAT + 11520
-----+-----+-----+-----+-----+-----+-----+
ACAGATAGCCATGATTTTGGGTATACACAAATTATGTTGTTAAGGTTCAATACGTTGTA
      S I G T K I P Y V V N T T I P S N A T F -
11521 TTGCAACTTCATTTTGGTCAGATGAATGACAGAGGTCTAACTTATATAATGAAGATGTAA + 11580
-----+-----+-----+-----+-----+-----+-----+
AACGTTGAAGTAAACCCAGCTCTACTTTACTGTCTTCCAGATTGAATATTACTTCTACATT
      A T S F W S D E M T E G L T Y N E D V T -
11581 CAATTACTTTGAATAATGTAGCTATGGATCAAGCTGATTATGAGTCACATAAGGAATA + 11640
-----+-----+-----+-----+-----+-----+-----+
GTTAATGAACCTTATTACATCGATACCTAGTTCGACTAATACTTCAGTGATTTCCTTTAT
      I T L N N V A M D Q A D Y E V T K G N N -
11641 ATGGCTTTAACTTAAATTAACAGACGAGGTTTAGCTAAATAATTAATGTAAGGATGCAG + 11700
-----+-----+-----+-----+-----+-----+-----+
TACCGAAATTGAATTTTAAATTGCTTCGTCCAAATCGATTTTAAATTACCATTCCTACGTC
```


Figure 101AB

C	G F N L K L T E A G L A K I N G K D A D -
11701	ACCAAAAATCCAAATTACTTACTCAGCTACTTTGAACTCACCTTGCTGTTCAGACATTC -----+-----+-----+-----+-----+ TGTTTTTTTAGTGTTAATGAATGAGTCGATGAAACCTGAGTGAACGACAACCTCTGTGAAG
C	Q K I Q I T Y S A T L N S L A V A D I P -
11761	CTGAAAGTAACGATATTACATATCATTAACGGAAATCATCAAGATCATGGGAATACTCCA -----+-----+-----+-----+-----+ GACITTCATTCGTATATAATGATAGTAATGCCITTATAGTTCCTAGTACCCTTATGAGTTT
C	E S N D I T Y H Y G N H Q D H G N T P K -
11821	AACCAACTAAACCTAATAATGGTCAAAFTACAGTAACATAAGACATGGGACAGTCAACCTG -----+-----+-----+-----+-----+ TTGGTTGATTGGATTATACCAGTTTAAATGTCATTGATTCCTGTACCTCTGCAGTTGGAC
C	P T K P N N G Q I T V T K T W D S Q P A -
11881	CTCCTGAGGGGTGAAAGCGACTGTTCAACTTGTAATGCCAAGACTGGTGAGAAAAGTCG -----+-----+-----+-----+-----+ GAGGACTCCCCACTTTTCGGCTGACAAGTTGAACATTACGGTTCTGTGACCACTCTTTTCAGC
C	P E G V K A T V Q L V N A K T G E K V G -
11941	GTGTCCTCTGTAGAACCTTTCAGAAAAATAATGGACATATACCTTGAGTGGTCTAGATAATT -----+-----+-----+-----+-----+ CAGGAGGACATCTTGAAAGTCITTTTATTAACTGTHATATGAACCTCACCAAGATCTATTAA
C	A P V E L S E N N W T Y T W S L D N S -
12001	CTATTGAATACAAAAGTTGAAGAAGATATAATGGATFACTCAGCTGAATACACAGTAGAGA -----+-----+-----+-----+-----+ GATAACTTATGTTTCAACTTCTTCTTATATACCTATGAGTCGACTTATGTGTCACTCT
C	I E Y K V E E E Y N G Y S A E Y T V E S -
12061	GCAAAGGGGAAGTTGGGGTAAAAAACCTGGAAAGATAATAACCCAGCTCCAATCAATCCTG -----+-----+-----+-----+-----+ -----+-----+-----+-----+-----+

280/487

Figure 101AC

```
CGTTCCCTTCAACCCCAATTTTGTGACCTTTCTATTATTTGGTTCGAGGTAGTAGGAC
      K G K L G V K N W K D N N P A P I N P E -
12121 -----+-----+-----+-----+-----+-----+-----+-----+
AAGAACACGCTGTAAACACATACGGTAAAGAGTTTGTCAAGTAGACCAAAAGATACCTC
      E P R V K T Y G K K F V K V D Q K D T R -
12180 -----+-----+-----+-----+-----+-----+-----+
TTCCTGGTGCACATTTTGTATGCCATTTTCAAAACAGTTTCATCTCGGTTTTTCTATGAG
      G T C T A G A A A T G C G C A G T T C G T T G T T A A A A A G C A G A T A C A A T A T A T T C C C T T T A
12240 -----+-----+-----+-----+-----+-----+-----+
CAGATCTTTTACGGCTCAGCAACAATTTTTCGTCATCGTTATTTATATATAACGGAAT
      L E N A Q F V V K K A D S N K Y I A F K -
12241 -----+-----+-----+-----+-----+-----+-----+
AGTCAACTGCACACAAAGCTGCAGATGAAGAAAGCAGCAGCAACTGCAAAACAAAAATTGG
      S T A Q Q A A D E K A A A T A K Q K L D -
12300 -----+-----+-----+-----+-----+-----+-----+
TCAGTTGACGTTGTGTCGACGCTACTTTTTCGTCGTCGTTGACGTTTGTGTTTAAACC
      A T G A G C G G T A G C A G C T T A C A A A T G C T G C A G A T A A G C C G C T C A A G C T A G T A G
12360 -----+-----+-----+-----+-----+-----+-----+
TACGTGGCCATCGTCGAATGTGTTACGACGCTCTATTTCGTTCCGGCGAGTTGAGATCATC
      A A V A A Y T N A A D K Q A A Q A L V D -
12361 -----+-----+-----+-----+-----+-----+-----+
ATCAAGCACAGCAAGAATACAAATGATAGCTTACAAAGAACCAAAATTTGGTTATGTTGAAG
      Q A Q Q E Y N V A Y K E A K F G Y V E V -
12420 -----+-----+-----+-----+-----+-----+-----+
TAGTTGCTGCTCTTATGTTTACATCGAATGTTTCTTCGGTTTAAACCAATACAACTTC
      T A G C T G G A A A A G A A G C A A T G G T T C T T A C T T C T A A T A F A C G G A T G F C A A T T C C A A A T T T
12480 -----+-----+-----+-----+-----+-----+-----+
ATCGACCTTTTCTACTTCGTTACCAAGAATGAAGATTATGCCTACGAGTTAAGGTTTAAA
      A G K D E A M V L T S N T D G Q F Q I S -
```

[illegible]

PCT/US05/27239 282/487

Figure 101AE

a V M A L C F S L V W G A H A V Q A Q E D -
CACACGTTGGTCTTGGCAATTGGAGAACTATCAGGAGGTGGTTAGTCAATTGCCATCTCGT 12901
GTGCGCAACCAGAACGTTAACTCTTGTAGTCTCTCCACCAATCAGTTAACGGTAGAGCA 12960
a H T L V L Q L E N Y Q E V V S Q L P S R -
GATGGTCATCGGTTGCAAGTATGGAAGTTGGATGATTCGTATTCTCTATGATCGGGTG 12961
CTACAGTAGCCCAACGTTTCATACCTTCAACCTACTAAGCATAAGGATACTACTAGCCCCAC 13020
a D G H R L Q V W K L D D S Y S Y D D R V -
CAAAATTGTAAGAGACTTGCCATTCTGGGATGAGAAATAAAGTTCTCTTCTTCAAAAAGACT 13021
GTTTAACATTCTCTGAACGTAAGCACCCCTACTCTTATTGAAAGAGAAAGTTTCTTGA
a Q I V R D L H S W D E N K L S S F K K T -
TCGTTGAGATGACCTTCCTTGAGATCAGATGGAATGATCTCATATCCAAATGGTCTT 13081
AGCAAACTCTACTGGAAGGAAGTCTTACTTACTTATGATAGAGATTAAGTTTACCAGAA 13140
a S F E M T F L E N Q I E V S H I P N G L -
TACTATGTTGGTCTATTATCCAGACGGATGCGGTTTCTTATCCAGCTGAATTCCTTTT 13141
ATGATACAAGCGAGATAATAGGTCTGCTTACGCCCAAGAATAGGTGACTTAAAGAAAAA
a Y Y V R S I I Q T D A V S Y P A E F L F -
GAAATGACAGATCAACGGTAGAGCCTTTGGTCAATTGTAGCGAAAAAACAGATACAATG 13201
CTTTACTGTCTAGTTTGCCATCTCGGAACCAGTAAATCGCTTTTGTCTATGTTAC 13260
a E M T D Q T V E P L V I V A K K T D T M -
ACACAAAGGTGAAGCTGATAAAGGTGGATCAAGACCAATCGTTGGAGGTGTCTGGC 13261
13320

PCT/US05/27239 283/487

Figure 101AF

13321 TGTGTTTCCACTTCGACTATTTCACCTAGTTCTGGTGTAGCGAACCTCCACAGCCG
a T T K V K L I K V D Q D H N R L E G V G -
TTTAAATTGGTATCAGTAGCAAGAGATGGTCTCTGAAAAAGAGGTTCCCTTGATTGGAGAA
13380 +-----+
AAATTAACCATAGTCATGGTTCTCTACCAAGACTTTTTCCTCAAGGGAACATAACCTCTT
a F K L V S V A R D G S E K E V P L I G E -
TACCGTTACAGTTCCTCTGGTCAAGTAGGGAGAGACTCTCTATACTATGATAAAAATGGAGAG
13440 +-----+
ATGGCAATGTCAGAGACCAAGTTCATCCCTCTCTGAGAGATGACTATTTTACCTCTC
a Y R Y S S S G Q V G R T L Y T D K N G E -
ATTTTGTGACAAATCTCTCTTGGGAACATATCGTTTCAAGGAGGTGGAGCCACTGGCA
13500 +-----+
TAAACACTGTTTAGAAGGAGAACCTTGATAGCAAGTCTCCACCTCGGTGACCGT
a I F V T N L P L G N Y R F K E V E P L A -
GGTATGCTGTACGACGCTGGATACGGATGTCAGCTGGTAGATCATCAGCTGGTGACG
13560 +-----+
CCGATACGACAATGCTGGACCTATGCTACAGTCCGACCATCTAGTAGTCGACCACTGC
a G Y A V T T L D T D V Q L V D H Q L V T -
ATTACGGTGTCAATCAGAAATTACCACGTTGGCAATGTTGACTTTATGAAGGTGGATGGT
13620 +-----+
TAATGCCAACAGTTAGTCTTTAATGGTGACCGTTACACTGAAATACTTCCACCTACCA
a I T V V N Q K L P R G N V D F M K V D G -
CGGACCAATACCTCTCTTCAAGGGCAATGTTCAAGTCAATGAAGAAGAAAGCGGACAC
13680 +-----+
GCCTGGTATGGAGAGAAGTTCCTCGTTACAGTTTTCAGTACTTTCTCTTTCGCTGTG
a R T N T S L Q G A M F K V M K E S G H -

Figure 101AG

13681	a	TATACTCCTGTTCTTCAAAATGGTAAAGGAGTAGTGTGTTAAACATCAGGGAAGAGATGGTCTGT -----+-----+-----+-----+-----+-----+-----+-----+ ATATGAGGACAAAGAAGTTTACCATTCTCCTTCATCAACATTTGTAGTCCCTTTCTTCTACCAAGCA	13740
	a	Y T P V L Q N G K E V V V T S G K D G R - -----+-----+-----+-----+-----+-----+-----+-----+ TTCCGAGTGGAAAGGTCTAGAGTATGGACATACATAATTTATGGGAGCTCCAAGCTCCAACCT 13741-----+-----+-----+-----+-----+-----+-----+-----+ AAGGCTCACCTTCCAGATCTCATACCTCTGTATGATAAATAACCTCGAGGTTCCGAGGTTGA	13800
	a	F R V E G L E Y G T Y Y L W E L Q A P T - -----+-----+-----+-----+-----+-----+-----+-----+ GGTTATGTTCAANTTAACATCGCTGTTCTCTTTACATCGGGAAGATACCTCGTAAGGAA 13801-----+-----+-----+-----+-----+-----+-----+-----+ CCAAATACAAGTTAATGTAGCGGACAAAGGAAATGTTAGCCCTTTCTATGAGCATTCCTT	13860
	a	G Y V Q L T S P V S F T I G K D T R K E - -----+-----+-----+-----+-----+-----+-----+-----+ CTGTTAACAGTGGTTAAATAACAAGCGACCAAGGATGTATGTCAGATACAGGGGAA 13861-----+-----+-----+-----+-----+-----+-----+-----+ GACCATTTGTCACCAATTTTATTTATGTTCTGGTGGTGCCTAACTACAGGTCTATGTCCCTT	13920
	a	L V T V V K N N K R P R I D V P D T G E - -----+-----+-----+-----+-----+-----+-----+-----+ GAAACCTTGTAATATCTTGATGCTTGTGTCATTTTGTGTTTGGAGTGGTTATTATCTT 13921-----+-----+-----+-----+-----+-----+-----+-----+ CTTTGGAACATATAGAACTACGAACAACCGTAAACACAAACCAATCACCATAATAGAA	13980
	a	E T L Y I L M L V A I L L F G S G Y Y L - -----+-----+-----+-----+-----+-----+-----+-----+ ACGAAAAAACCAATAACTGATATTCATATGTACATCATATTATGAAAAAGATAGCAGGCTGA 13981-----+-----+-----+-----+-----+-----+-----+-----+ TGCCTTTTGTGGTTTATTGACTATAAGTTACATGTAGTAAATACCTTTTCTATCGTCCGACT	14040
	a	T K K P N N * -----+-----+-----+-----+-----+-----+-----+-----+ AGGGAAGACCAGAGTACTCTCAGGTGATGTTAATCAGGAATCATGGTGATGGGCATGAA 14041-----+-----+-----+-----+-----+-----+-----+-----+ TCCCTTCTGGTCTCATGAGACTCCACTCAATTAGTCTCTTAGTACCCTACACCGTACTT	14100

Figure 101AH

14101 TCACAATAACGGATATAGGCTGGCAGATGTGCCAGCCTCATGTGGGTATTGTTTG + 14160
+-----+-----+-----+-----+-----+
AGTGTATTGCTATACCTCGACCCGCTCTAACACGGTGGAGTAACACCCCAATAACAAAC
14161 TAAAAAGATAGGACTGGTCTGGTAATCATTTTAGGAATGGACAGACTGGGATTCGTATT + 14220
+-----+-----+-----+-----+-----+
ATTTTGTATCCTGACCCAGACCATTAAGTAAATCCTTACCTGCTCCTGACCCCTAAGACTAA
14221 TAAATGGATGGTGAATCAGAAAGAAATGAGATTTTCTCGTTTCTCTTAGCAGATAGGAT + 14280
+-----+-----+-----+-----+-----+
ATTTTACCTACCACTTAGTCTTCTTTTACTCTCTAAAGAGCAAGAGAAATCCTCTATCCTA
14281 TGTCTGTTAGGAAAAGCGATATAAATGATGAGTTTGAAGATAAAGGGATGCTGATATAAAAT + 14340
+-----+-----+-----+-----+-----+
ACAGACAAATCCTTTTTCGCTATTTTACTACTCAAACTCTATTTTCCCTACGACTATTTTAA
M L I K M -orf6_670, homologue of sp0466, sortase
CGTAAAAAGAAAAAGCAAAAACGAAATAATCTCCTATTAGGAGTGGTATTTTCATTGG + 14400
+-----+-----+-----+-----+-----+
CCATTTTGTGTTTTTTCGTTTTTTCCTTTATTAGAGGATAATCCTCACCATAAAAAGTAACC
V K T K K Q K R N N L L L G V V F F I G -
14401 AATGGCGTAATGGCGTATCCGCTGGTGTCTCGTCTGTTATTATCGAGTGGATCAAAATCA + 14460
+-----+-----+-----+-----+-----+
TTACCGCCATACCGCATAGCGGACCCACAGAGCGGAACATAATAGCTCACCCTAGTTTAGT
M A V M A Y P L V S R L Y Y R V E S N Q -
14461 ACAAAATTGCTACTTTGATAAGGAAAAACCAACGTTGGATGAGGCTGACATGATGAACG + 14520
+-----+-----+-----+-----+-----+
TGTTTAACGACTGAACCTATTCCTTTTTCGTTGTCACCTACTCCGACTGTAACTACTTGC
Q I A D F D K E K A T L D E A D I D E R -
AATGAAATGGCACAAGCCTTCAATGACTCTTTTGAATAAATGATGAGTGGCGATCCTTG

[illegible]

14941	b	GATTGTACCAGGTCATGATTATGTGACCTTGCTGACTTGTACGCCATACATGATCAATAC -----+-----+-----+-----+-----+-----+-----+ CTAACATGGTCCAGTACTAATAATACACTGGAAAGACTGAACATCGCGGTATGTACTAGTTATG -----+-----+-----+-----+-----+-----+-----+ I V P G H D Y V T L T C T P Y M I N T - -----+-----+-----+-----+-----+-----+-----+ CCATCGTCTATTTGGTTCGGGGGCATCGGATACCGTACGTAGCAGAGTTGAGGAAGAAATT -----+-----+-----+-----+-----+-----+-----+ GGTAGCAGATAACCAAGCCCCCGTAGCCTATGGCATGCATCGTCTCCCAACTCCTTCTTAA -----+-----+-----+-----+-----+-----+-----+ H R L L V R G H R I P Y V A E V E E F - -----+-----+-----+-----+-----+-----+-----+ TATTGCAGCAAAACAACTCAGTCATCTCTATCGCTACCTGTTTATGTGGCAGTTGGTTTT -----+-----+-----+-----+-----+-----+-----+ ATAACGTCGTTTGTGTCAGTCAGTAGAGATAGCGATGGACAAATAFACACCGTCAACCCAAA -----+-----+-----+-----+-----+-----+-----+ I A A N K L S H L Y R Y L F Y V A V G L - -----+-----+-----+-----+-----+-----+-----+ GATTGTGATTCCTTTTATGGATTATTCGACGCTTGGCAAGAAAGAAAAACAACCGGAAAA -----+-----+-----+-----+-----+-----+-----+ CTAACACTAAGAAAAATACCTAATAAGCTGCGAAGCGCTTCTCTTTTGTGTTGGCCTTTT -----+-----+-----+-----+-----+-----+-----+ I V I L L W I I R R L R K K K Q P E K - -----+-----+-----+-----+-----+-----+-----+ GGCTTTGAAGCGCGCTGAAAGCAGCAAGGAAGGAAGTGAAGGTGGAGGATGGACAACAGTA -----+-----+-----+-----+-----+-----+-----+ CCGAAACTTCCGGGACTTTCGTCGTTCCCTTCCTTCCACTTCCACCTCCTACCTGTGTCAT -----+-----+-----+-----+-----+-----+-----+ A L K A L K A A R K E V K V E D G Q Q * - -----+-----+-----+-----+-----+-----+-----+ GACGTTACGAAAAAAGGCACAAAAAAGAAACATCCCGCTGATCCTCTCTCTGATTT -----+-----+-----+-----+-----+-----+-----+ CTGCAAGTCCTTTTTCGGTGTCTTTTCTCTCTTTGTAGCGGACTAGGAAGAGACTAAA -----+-----+-----+-----+-----+-----+-----+ TCTTAGTAGGATTCCGCGTTGGGATATATCCATTTGGTGTCTCGTTATTTATTCGTATTG -----+-----+-----+-----+-----+-----+-----+ AGAAATCATCCTTAAGCGGCAAGCGTATATAGGTAACCCACAGAGCAATAATAAAGACATAAC -----+-----+-----+-----+-----+-----+-----+ 15001 15061 15121 15181 15241 15301 15360
-------	---	---

VSRYRYRIE-orf7_670, homologue of sp0467, sortase

PCT/US05/27239

288/487

Figure 101AK

```
15361 AGTCAAACGAGGTTATTAAAGAGTTTGATGAGACGGTTTCCAGATGGATAAGGCAGAAC 15420
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCAGTTTGCTCCCAATAATTTCCTCAAACTACTCTGCGCAAGGGTCTACCTATTCCTCGTCTTTG
C S N E V I K E F D E T V S Q M D K A E L -
15421 TTGAGGAGCGTTGGCGCTTGCTCAAGCCTCAATGCGACCTTGAACCATCTGAAATTC 15480
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACTCCTCGCAACCGCGAACCGAGTTCCGGAAGTTACGCTGGAACCTTTGGTAGACTTTTAAG
C E E R W R L A Q A F N A T L K P S E I L -
15481 TTGATCCTTTTACAGAGCAAGAGAAAAAGAAAGCGCTCTCAGAAATATGCCAATATGCTAA 15540
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACTAGGAAAAATGTCTCGTTCGTTCTCTTTCTTTTCCGCGAGAGTCTTATACGGTTATACGATT
C D P F T E Q E K K K G V S E Y A N M L K -
15541 AGTCCATGAGCGGATTGGCTATGTGGAATTCCTCGGATTGATCAGGAAATTCGGATGT 15600
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCCAGGTACTCGCCTAACCGATACACCTTTAAGGACGCTAACTAGTCTCTTTAAGGCTACA
C V H E R I G Y V E I P A I D Q E I P M Y -
15601 ATGTCCGAACGAGTGAGGAAATTCCTTCAGAAAGGCGCGAGGATTGCTAGAGGGAGCTTCGT 15660
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACAGCCTTGCTCAGTCTCTTTAAGAAGTCTCCCGGCTTAACGATCTCCCTCGAAGCA
C V G T S E E I L Q K G A G L L E G A S L -
15661 TACCGGTTGGTGGTGAATAATACCCACACAGTTGTCTCAGTCTCATAGAGGATTACCGACGG 15720
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGGCCAACCAACCACTTTTATGGGTGTGTCAACAGTGACGAGTATCTCCTAATGGCTGCC
C P V G G E N T H T V V T A H R G L P T A -
15721 CAGAACTGTTTAGTCAATTGGATAAGATGAAAAAGGGGATGCTTTTATCTTCACGTTT 15780
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTCCTTGACAAATCAGTTAACCTATTCTACTTTTTTCCCTTACAGAAAAATAGAAAGTGCAAA
```

PCT/US05/27239

289/487

Figure 101A

```
c      E L F S Q L D K K K G D V F Y L H V L -
15781  TAGACCAGGTGTTGGCCTACCAAGTGGATCAGATTTTGACGGTTGAGCCAAATGACTTTG + 15840
      ATCTGGTCACAACCGGATGGTTCACCTAGTCTATAAAGCTGCCAACTCGGTTTACTGAAAC
c      D Q V L A Y Q V D Q I L T V E P N D F E -
15841  AGCCTGTCTTGATTCAACATGGGGAGATTATGCGACCTTGTGTGACCTGTACACCGTATA + 15900
      TCGGACAGAACTAAGTTGTACCCCTCTATAAGCTGGACAACTGGACATGTGGCATAT
c      P V L I Q H G E D Y A T L L T C T P Y M -
15901  TGATTAAAGTCAATCGTCTGTGTTGGTACGTTGGGAGCGGATTCCGTATACGGCACCAATTG + 15960
      ACTAATTGTCAGTAGCAGACAAACCATGCACCTTCCCTAGGCATATGCCGTGGTTAAC
c      I N S H R L L V R G K R I P Y T A P I A -
15961  CAGAGCGAATCGAGCGGTGAGAGCGGTGGGAAATTCGTGTTGTTGTTATTGCTAGCGG + 16020
      GTCTCGCTTAGCTCGGCACCTCTCTGACCCGCTAAGACCAACACCAATACGATCGCC
c      E R N R A V R E R G Q F W L L L L A A -
16021  CGTTGGTTATGATTCTGGTATTGAGTTACGGGGGTATCGTCATCGTCCGATTGTCAAAG + 16080
      GCAACCAATACTAAGACATAACTCAATGCCCCACATAGCAGTAGCAGCGTAACAGTTTC
c      L V M I L V L S Y G V Y R H R R I V K G -
16081  GGCTAGAAAACAAATTGGAGGAGCATCATGTCAAGGCTAAGCTACAGAAATTACTAGGG + 16140
      CCGATCTTTTGTAACTCCTCGTAGTACAGTTCCGATTCGATGTCCTTAATGATCCC
a      M S K A K L Q K L L G -orf8_670, homologue of sp0468, sortase
c      L E K Q L E E H H V K G *
```

290/487

Figure 101AM

16141 TATTGCTGATGCTGGTAGCATTTGGTAGTTCCTGTTTATTTGTTTGGGAGATGGTGTTA 16200
-----+-----+-----+-----+-----+-----+-----+
ATAAACGACTAGACCATCGTAACCACTAAGGACAATAACAAAACCCGCTCTACCACAAT
a Y L L M L V A L V I P V Y C F G Q M V L -

16201 CAGTCTTTAGGACAAAGTAAAGTCATGAGATATTTTCAGAAATCTGTGACGGCCGACAGT 16260
-----+-----+-----+-----+-----+-----+-----+
GTCAGAAATCCTGTTTCATTTTCCAGTACTCTATAAAAAGTCTTAGACACATGCCCGGCTGTCA
a Q S L G Q V K G H E I F S E S V T A D S -

16261 TACCAAGACCAATTGCAACGGTCGCTTGATTACAAATCAACGCTGGATTGCGCAAAATCGT 16320
-----+-----+-----+-----+-----+-----+-----+
ATGGTTCCTGTTAACGTTGCCAGCGAACTAATGTTAGTTGCGAACCTAAGCGTTTATAGCA
a Y Q E Q L Q R S L D Y N Q R L D S Q N R -

16321 ATTGTAGATCCCTTTTGGCGGAAGGTATGAGTAAATACCAAGTCTCTGACGATCCT 16380
-----+-----+-----+-----+-----+-----+-----+
TAACATCTAGGAAAAAACCCGCTTCCCATCTCCATTTAATGGTTTCACAGACTGCTAGGA
a I V D P F L A E G Y E V N Y Q V S D D P -

16381 GATGCAGTCTACGGCTATTTGTCGATTCGAGTTGGAATCATGAGCCAGTTTATCTA 16440
-----+-----+-----+-----+-----+-----+-----+
CTACGTCAGATGCCGATAACAGCTAAGGCTCAACCTTTAGTACCTTCGGTCAAAATAGAT
a D A V Y G Y L S I P S L E I M E P V Y L -

16441 GGAGCGGATTACCATCATTTAGCAATGGGTTGGCCCATGTTGGATGGACGCCCTTCCT 16500
-----+-----+-----+-----+-----+-----+-----+
CCTCGCTAATGTTAGTAATTCGTTACCCCAACCGGGTACACCTACCTGCGGAGAAAGGA
a G A D Y H H L A M G L A H V D G T P L P -

16501 GTTGAGGGAAGGATTCGTTTCACTGATTGTTGGGACCGCTGCAGAACCAAGCCATGTC 16560
-----+-----+-----+-----+-----+-----+-----+
CAACTCCCTTTTCCCTAAGCAAGTCACTAACGACCCCGTGGCAGCTCTTGGTTCCGGTACAG

Figure 101AM

a V E G K G I R S V I A G H R A E P S H V -
16561 TTTTTCGGCCATTGGATCAGCTAAAGTTGGAGATGCTCTTTATATGATAATGGCCAG
-----+-----+-----+-----+-----+
AAAAGCGGTAAACCTAGTCGATTTTCAACCTCTACGAGAAATACTATTACCGGTC
16620
a F F R H L D Q L K V G D A L Y Y D N G Q -
16621 GAAATTGTAGAAATATCAGATGATGGACACAGAGATTATTTACCGTCGGAATGGGAAAA
-----+-----+-----+-----+-----+
CTTTAACATCTTATAGTCTACTACCTGTGTCTCTAATAAAATGGCAGCCTTACCCCTTTT
16680
a E I V E Y Q M M D T E I I L P S E W E K -
16681 TTGAATCGGTTAGCTCTAAAAATATCATGACCTTGATAACCTGGATCCGATTCCCTACC
-----+-----+-----+-----+-----+
AATCTAGCCAAATCGAGATTTTATAGTACTGGAACHTATTGGACGCTAGGCTAAGGATGG
16740
a L E S V S S K N I M T L I T C D P I P T -
16741 TTTAATAAACGCTTATTAGTGAATTTTGAACGAGTCGCTTTTATCAAAAAATCAGATCCA
-----+-----+-----+-----+-----+
AAATTATTGGGAATAATCACTTAAACCTTCTCAGCAGACAAATAGTTTATTAGTCTAGGT
16800
a F N K R L L V N F E R V A V Y Q K S D P -
16801 CAAACAGCTGCAGTTCGAGGGTTGCTTTACGAAAGAGGACAACTGTATCGCGTGT
-----+-----+-----+-----+-----+
GTTTGTGACGTCACGCTCCCAACGAAATGCTTTCTCTGTTAGACATAGCGCACAA
16860
a Q T A A V A R V A F T K E G Q S V S R V -
16861 GCAACTCTCAATGGTTGTACCGTGGGCTAGTGGTACTGGCATTTCTGGGAATCCTGTT
-----+-----+-----+-----+-----+
CGTTGGAGAGTTACCAACATGGCAGCCGATCACCATTGACCGTAAAGACCCCTTAGGACAAA
16920
a A T S Q W L Y R G L V V L A F L G I L F -
16921 GTTTTGTGGAAGCTAGCACGTTTACTACGAGGGAATAAAAAAGAAATGAAGGAAAGCTA
-----+-----+-----+-----+-----+
16980

PCT/US05/27239 292/487

Figure 101AN

a . V L W K L A R L L R G K *
CAAACACCTTCGATCGTGCAAAATGATGCTCCCTTTATTTTCTTTACTTTCCCTTTTCGAT
16981 AGGCTGTTTCCTTTTTCGGGCTCTTTTGTCAACTGTAGTGGGTGAAAAAGCTAAGCTCG
TCCGACAAGGAAAAAGGCGGAGAAACAGTTGACATCACCCAACTTTTTCGATTCGAGC 17040
17041 AGAAAGGACAAAATTTTGTCTCTTTCTTTTGTGATATTCAGAGCGATAAAAATCCGTTTTTT
TCTTTCCTGTTTAAACAGGAAGAAAAACATAAGTCTCGCTATTTTATAGGCAAAAAA 17100
GAAGTTTTCAAA
17101 -----+--- 17112
CTTCAAAAGTTT

M1, strain 2580

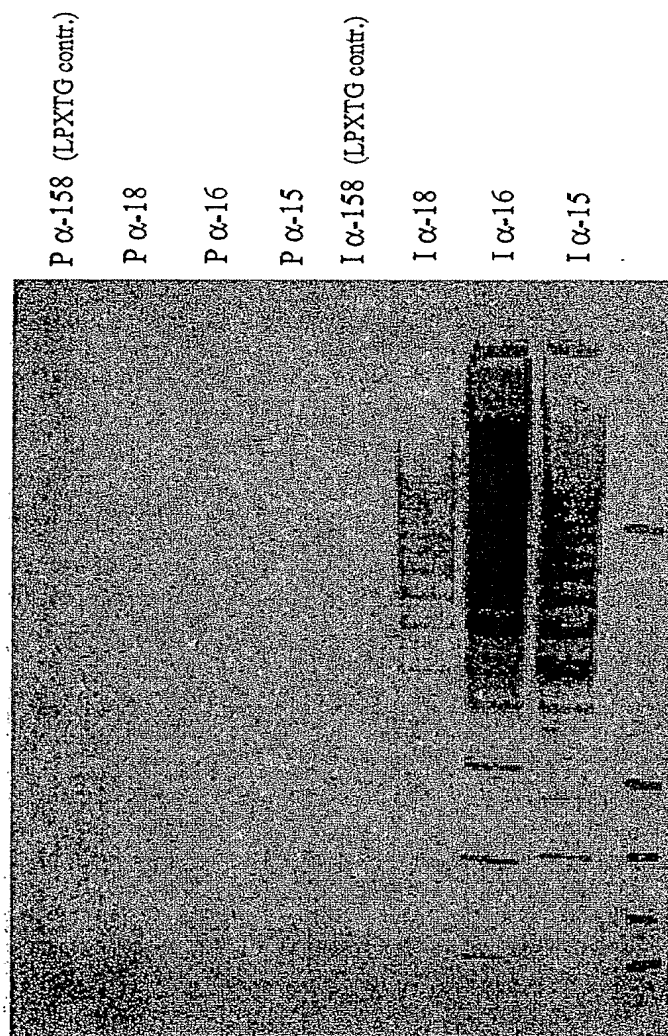


Figure 102

LEGEND:I α -#: immune serum anti-#P α -#: pre-immune serum anti-#

PCT/US2005/027239 294/487

M1, strain 2913

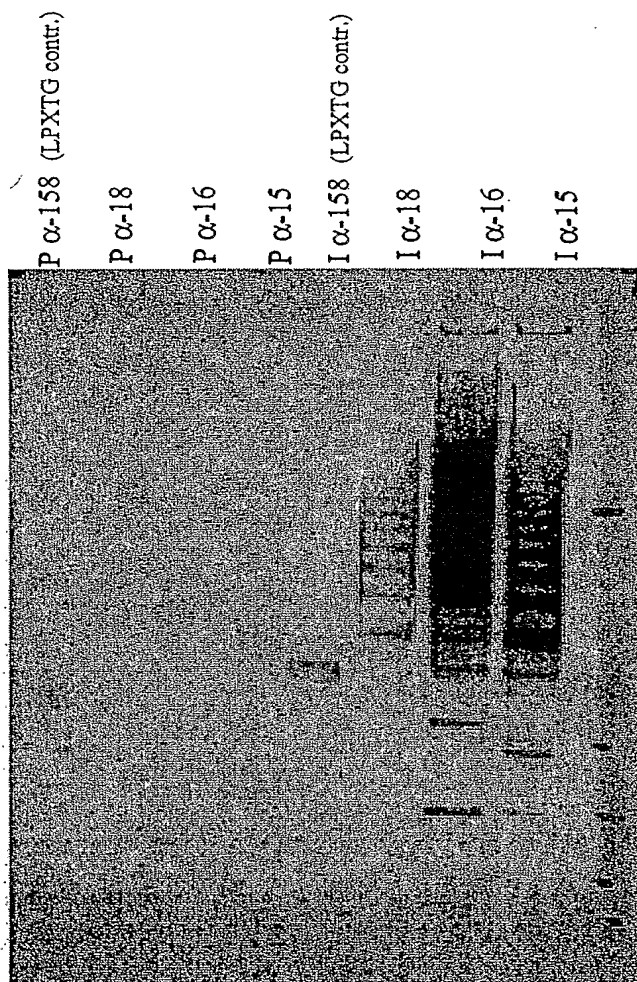


Figure 103

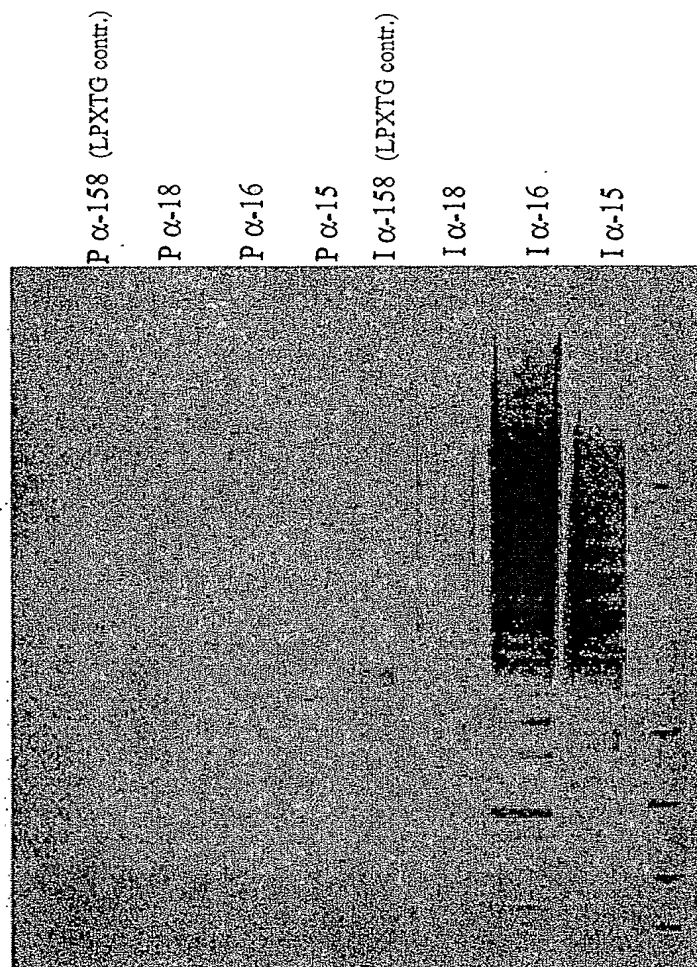
LEGEND:

I α-#: immune serum anti-#

P α-#: pre-immune serum anti-#

295/487

M1, strain 3280

**LEGEND:**

I α-#: immune serum anti-#

P α-#: pre-immune serum anti-#

Figure 104

296/487

M1 strain 3348

P α -158 (LPXTG contr.)
P α -18
P α -16
P α -15
I α -158 (LPXTG contr.)
I α -18
I α -16
I α -15

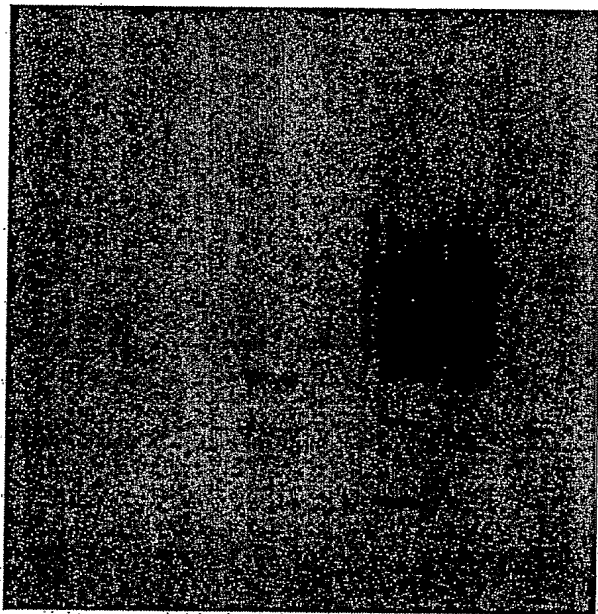


Figure 105

M1 strain 2719

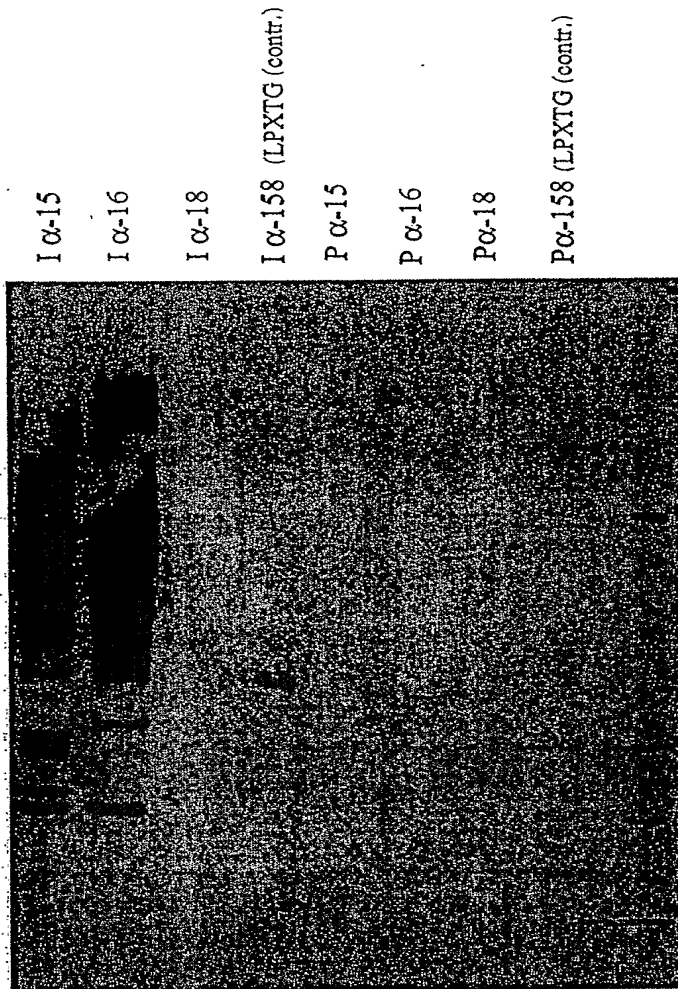


Figure 106

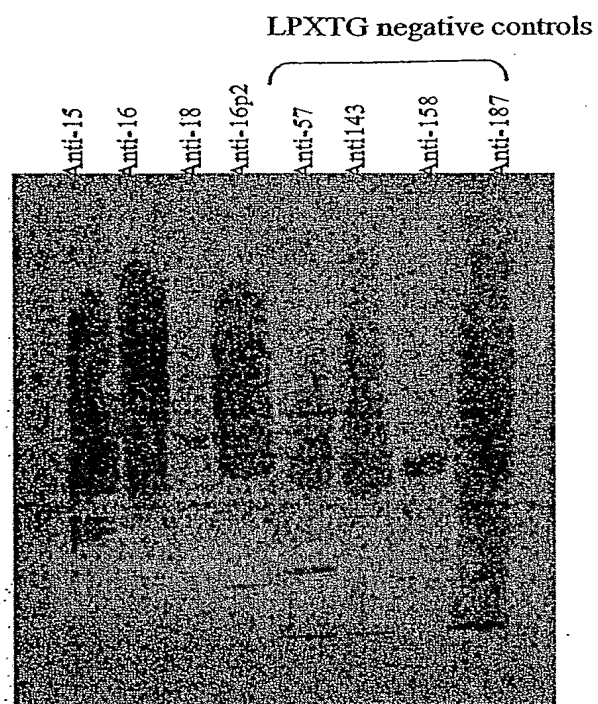
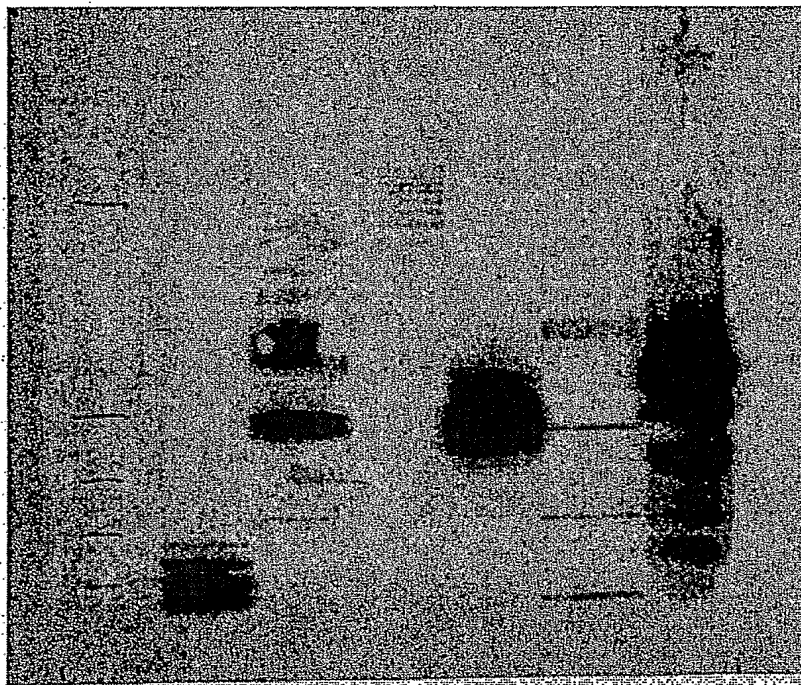
Western blot Western blot on fraction enriched in surface proteins of M1 (SF370)

Figure 107

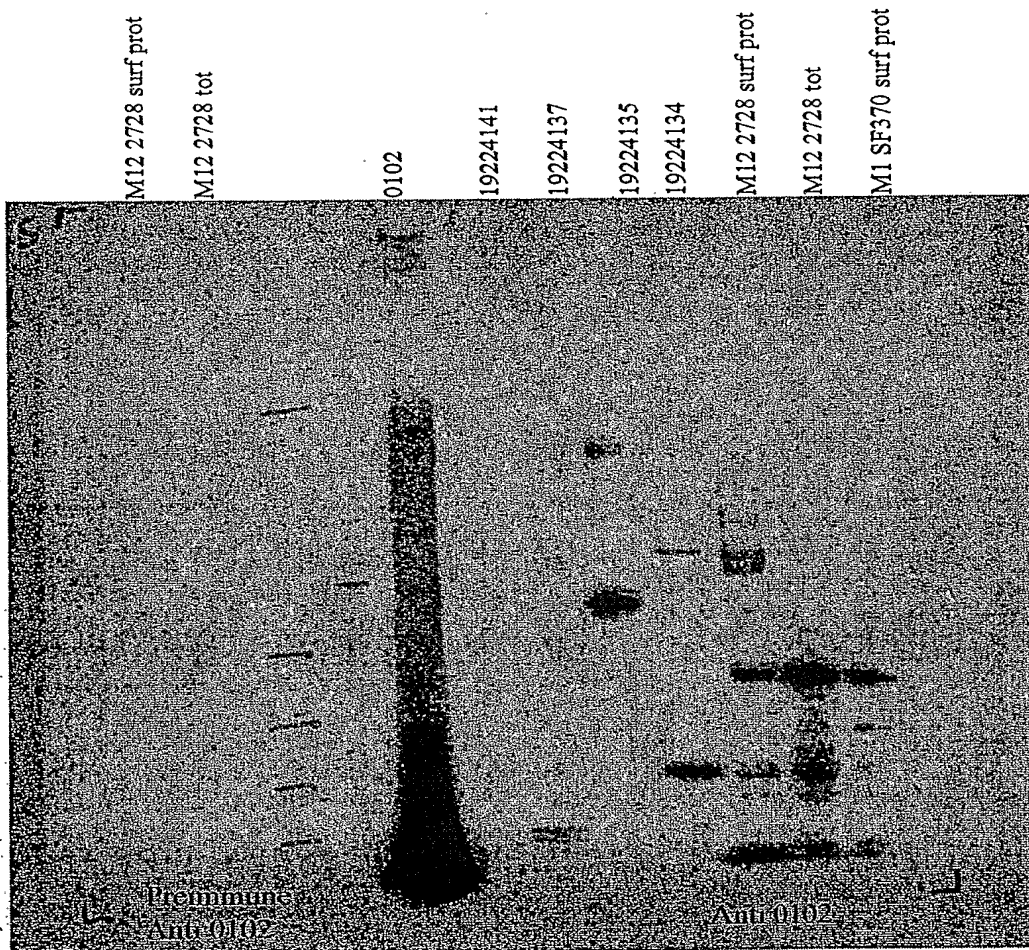
Western blot on fraction enriched in surface proteins of M12 (2728)

Anti-19224134
Anti-19224135
Anti-19224137
Anti-19224141
Anti-0102
Anti-158 (control)



Note: 0102 corresponds to SpyM3_0102 in MGAS315 (M3). This protein is pretty much identical to the M12 19224139 -see also gel "WB M12 (23-11-4b2a)"-, the M18 spyM18_0130 and the M5 orf 82. Here the antiserum raised against 0102 recognizes high-weight molecular structures in M12 (strain isolate 2728). This means that in M12 three out of the five LPKTG proteins form high-molecular weight structures:

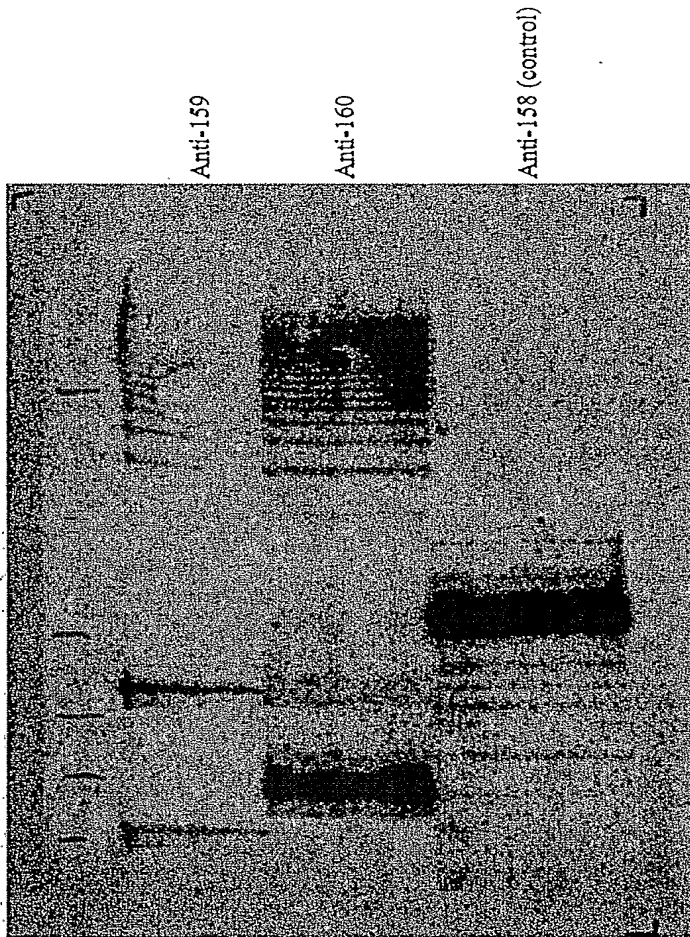
Figure 108



Note: 0102 corresponds to SpyM3_0102 in MGAS315 (M3). This protein is pretty much identical to the M12 19224139 -see also gel "WB M12 (23-11-4b2bis)"-, the M18 spyM18_0130 and the M5 orf 82. Here the antiserum raised against 0102 recognizes high-weight molecular structures in protein extracts from M12 (strain isolate 2728.)

Figure 109

Western blot on fraction enriched in surface proteins of M6 (2724)

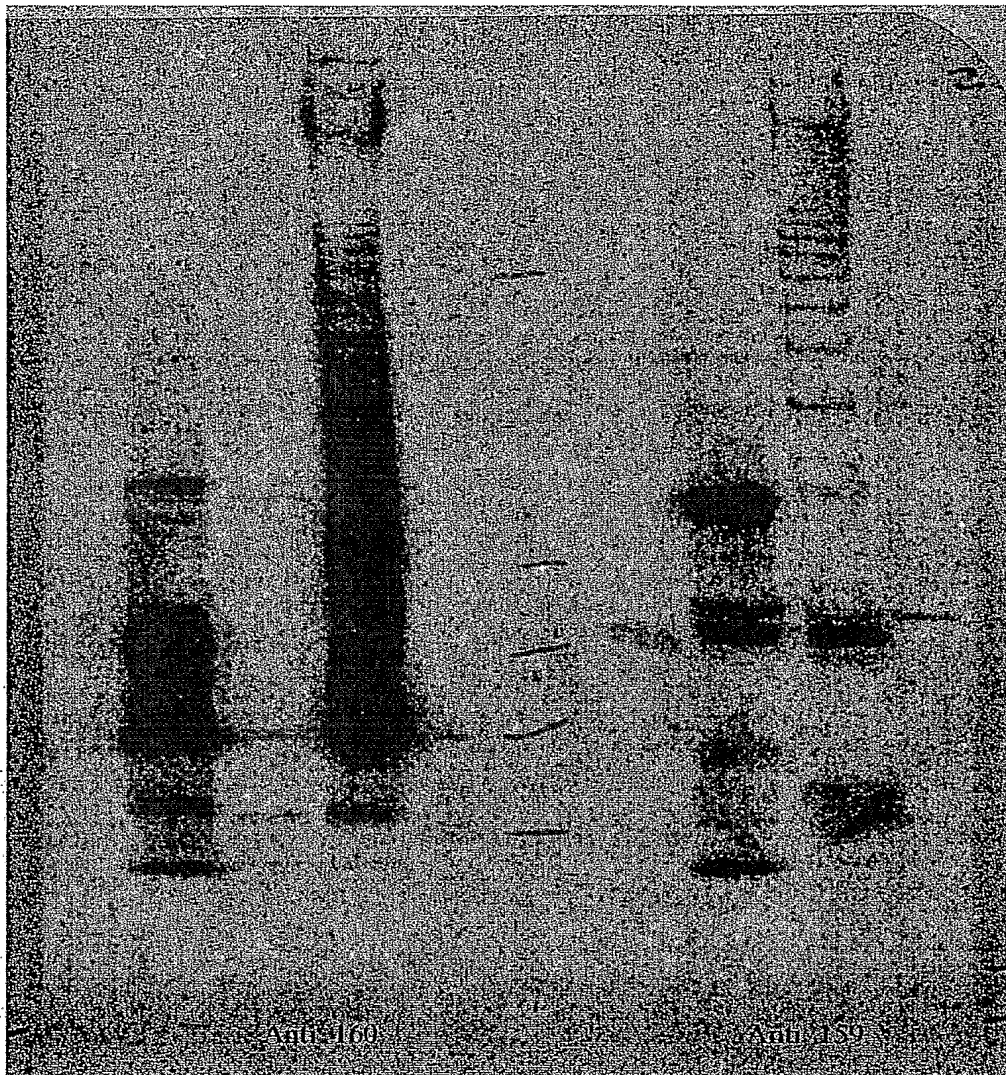


M6 strain isolate 2724

Figure 110

PCT/US05/27239

0160	0159	M6 3650 surf prot	M1 SF370 surf prot	Marker	0160	0159	M6 3650 surf prot	M1 SF370 surf prot
------	------	-------------------	--------------------	--------	------	------	-------------------	--------------------



M6 strain isolate 3650

Figure 111

PCT/US05/27239

M6 2724 fraz art

M6 2724 tot

0160

Marker

158

0160

0159

M6 2724 surf prot

M6 2724 tot

M1 SF370 surf prot

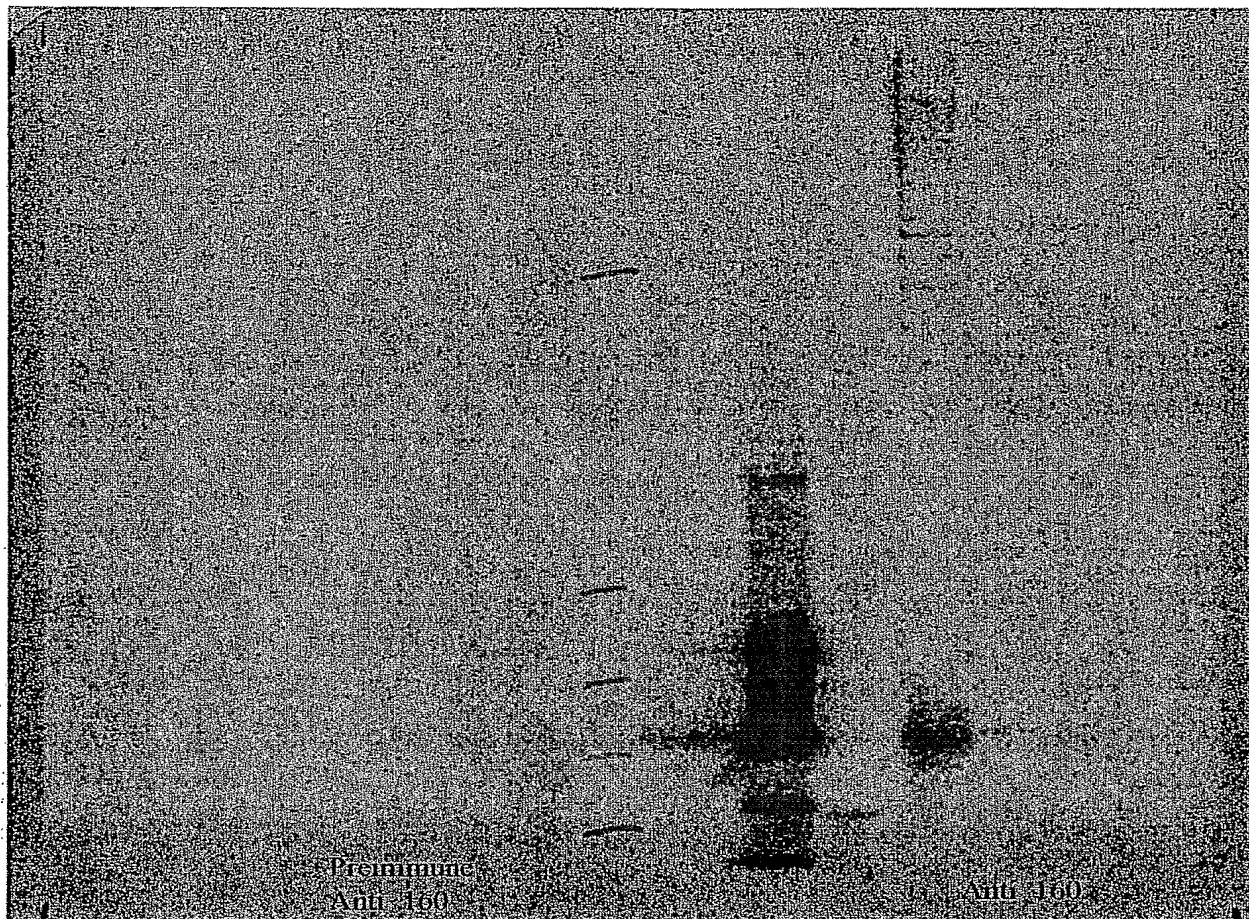


Figure 112

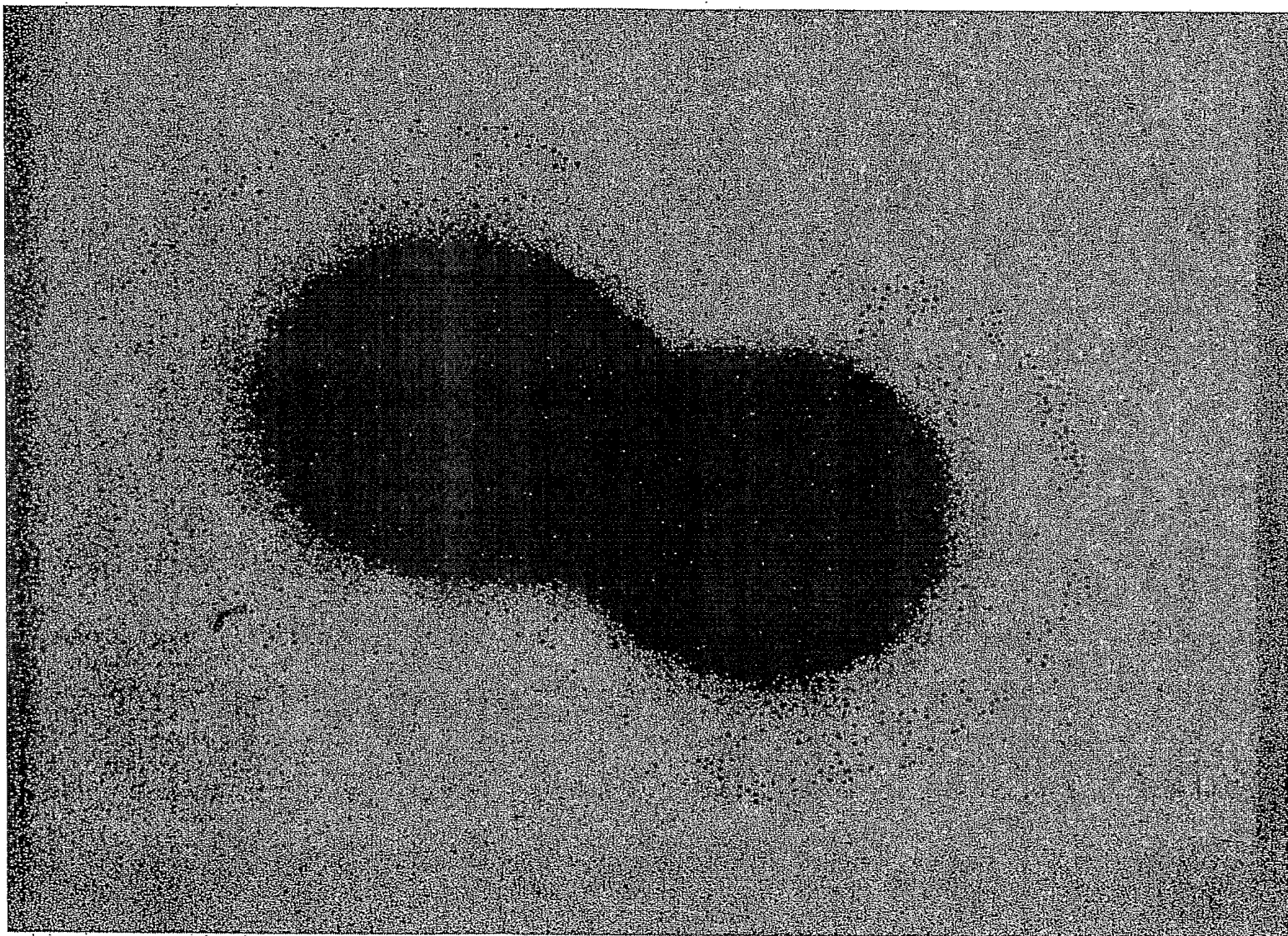


FIGURE 113

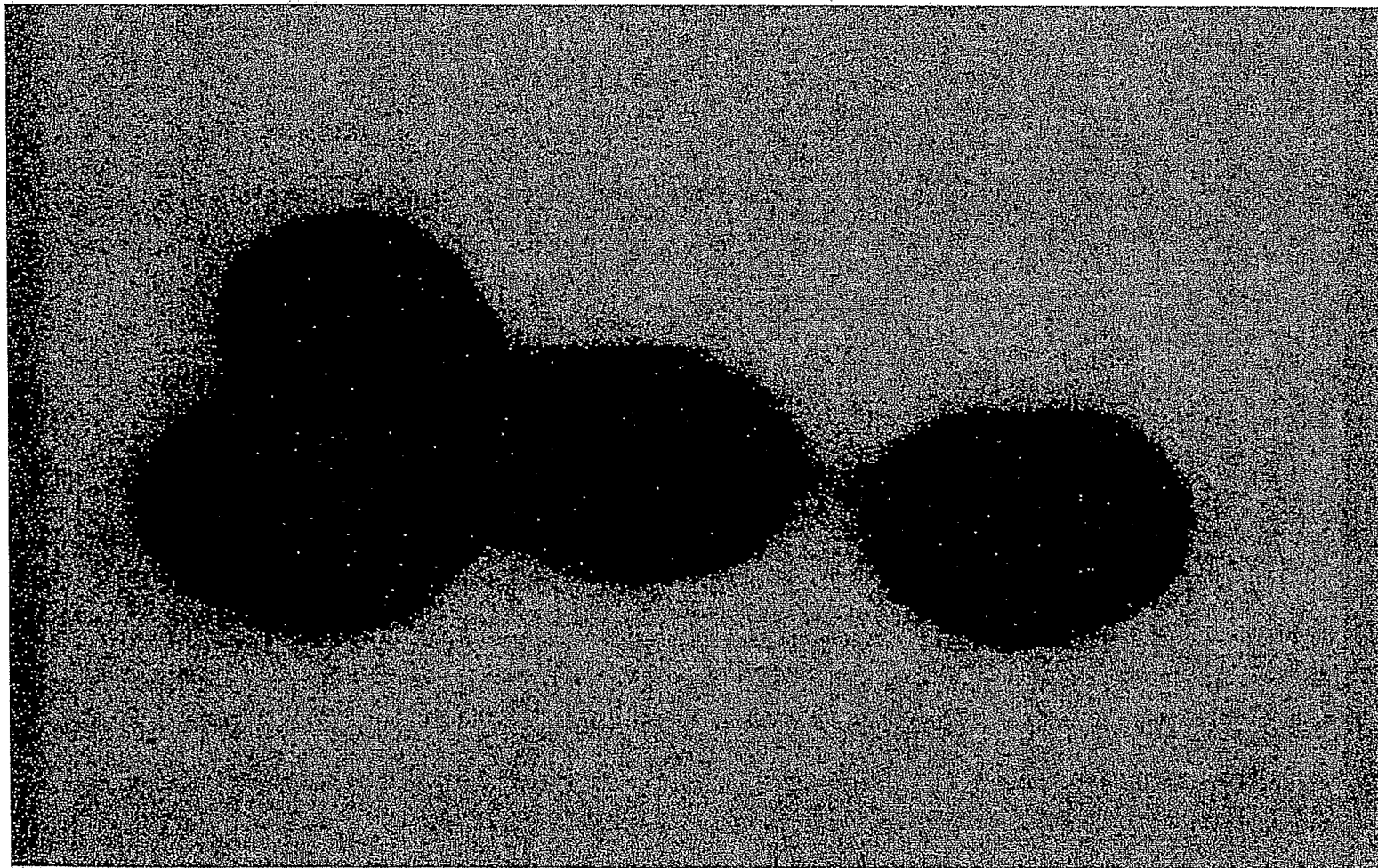


FIGURE 114

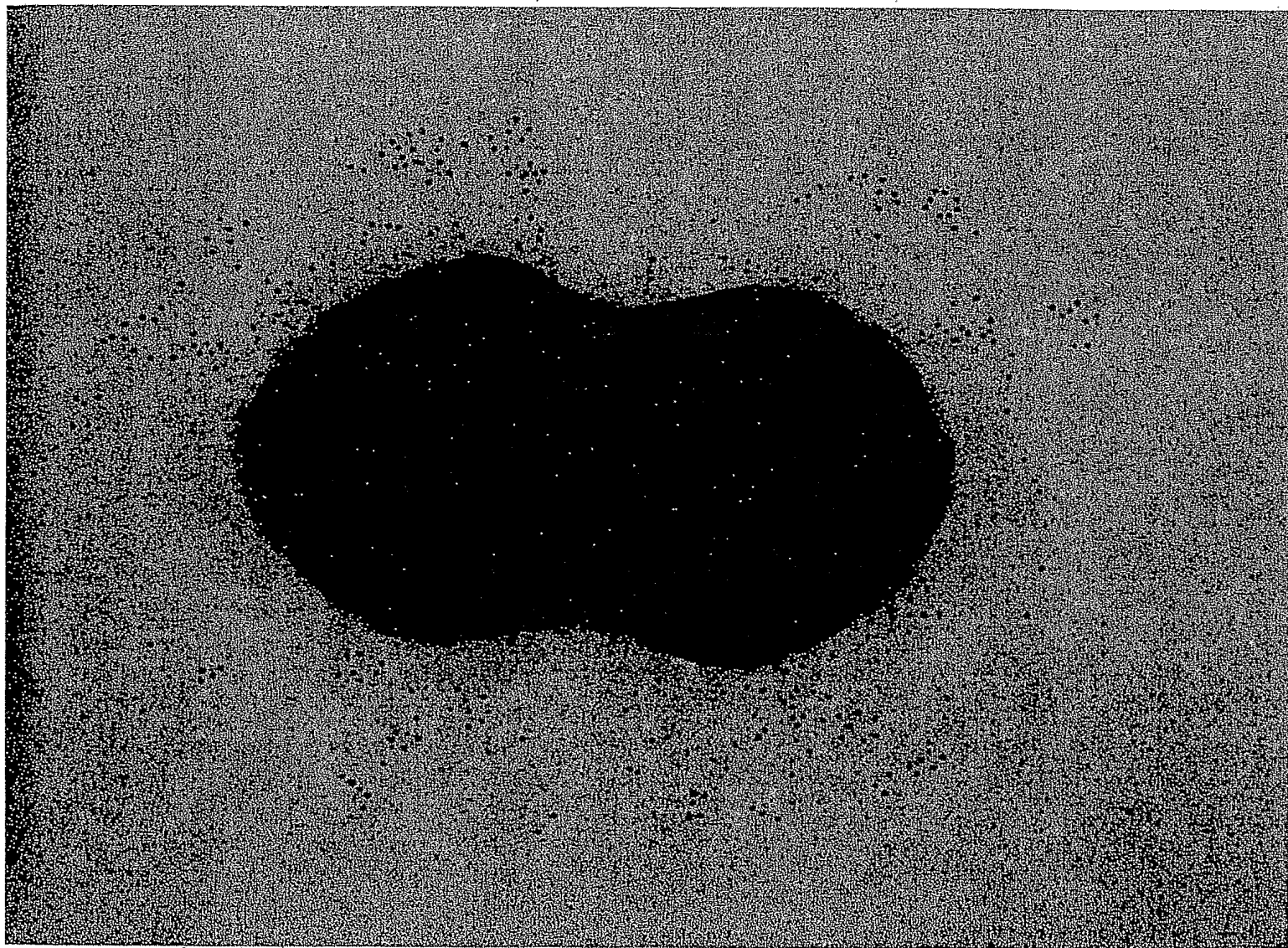


FIGURE 115

PCT/US05/27239 307/487

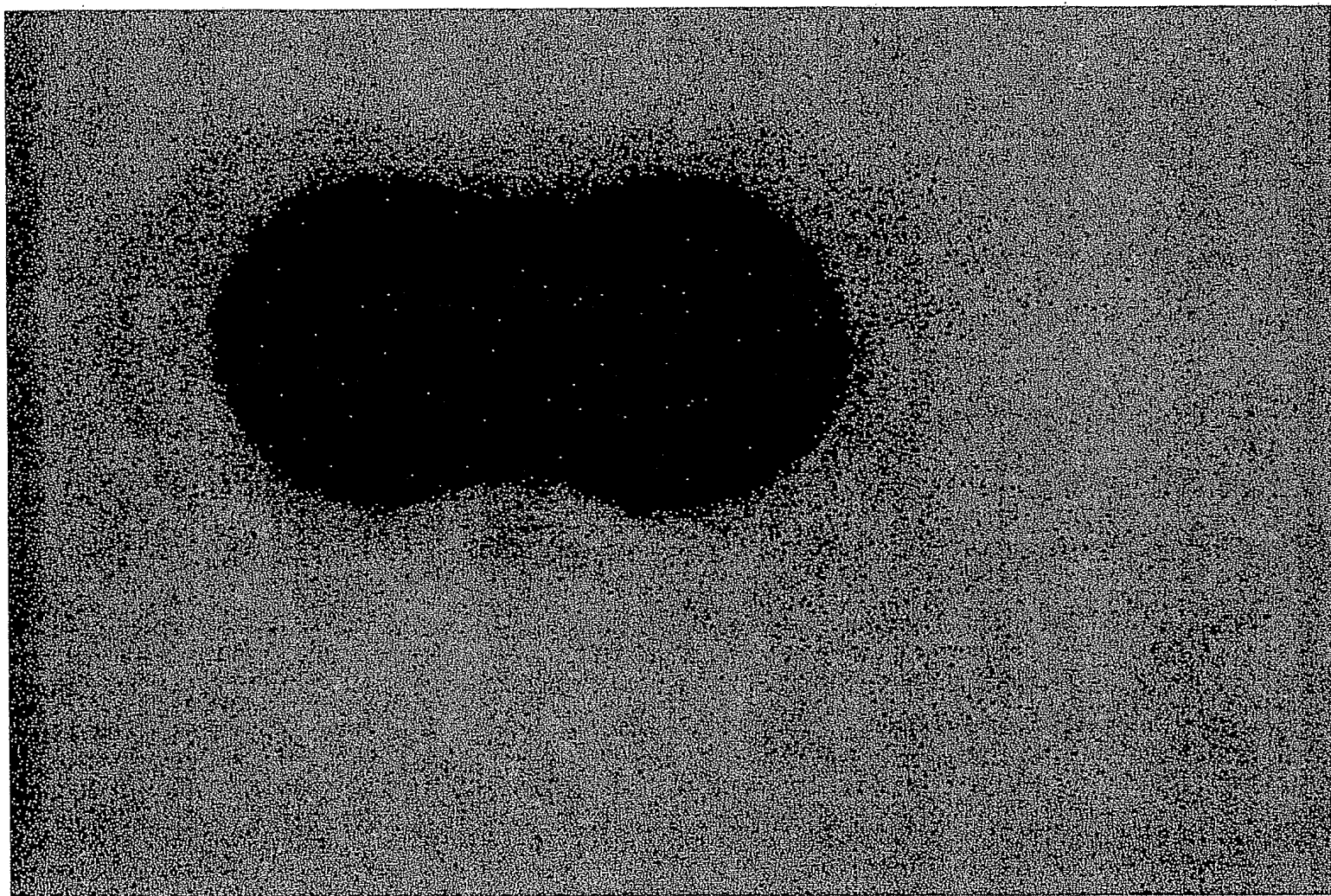


FIGURE 116

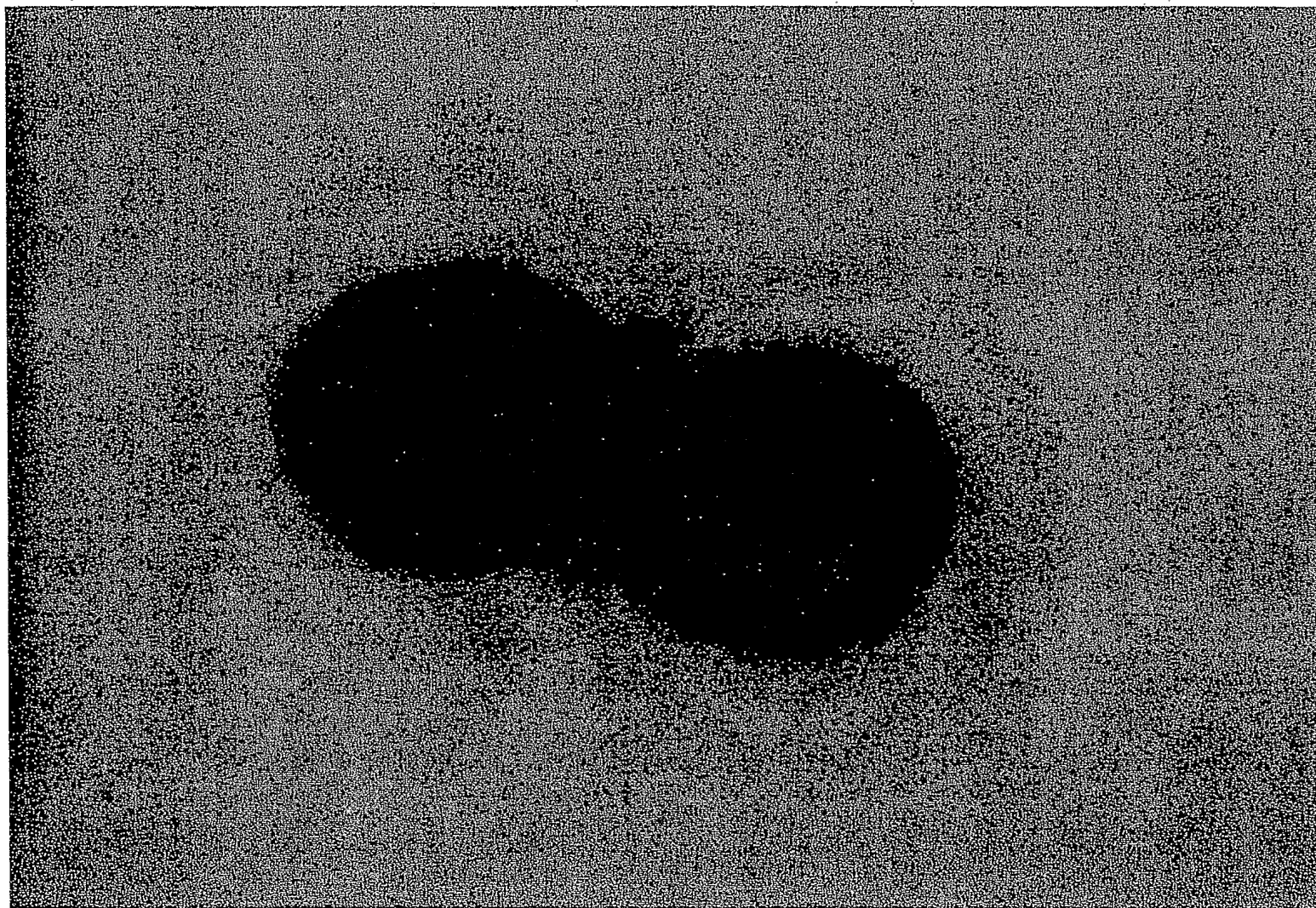


FIGURE 117

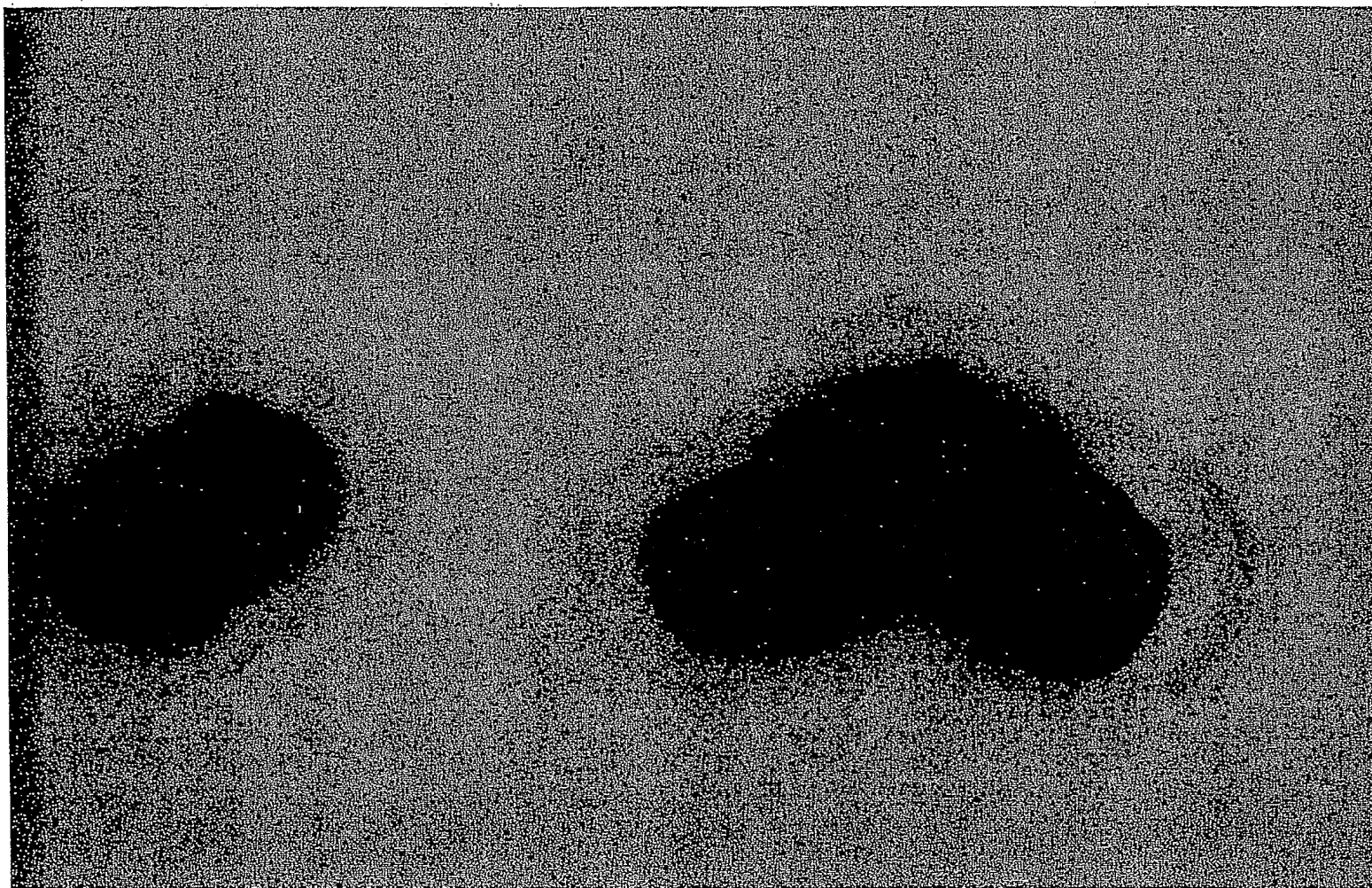


FIGURE 118

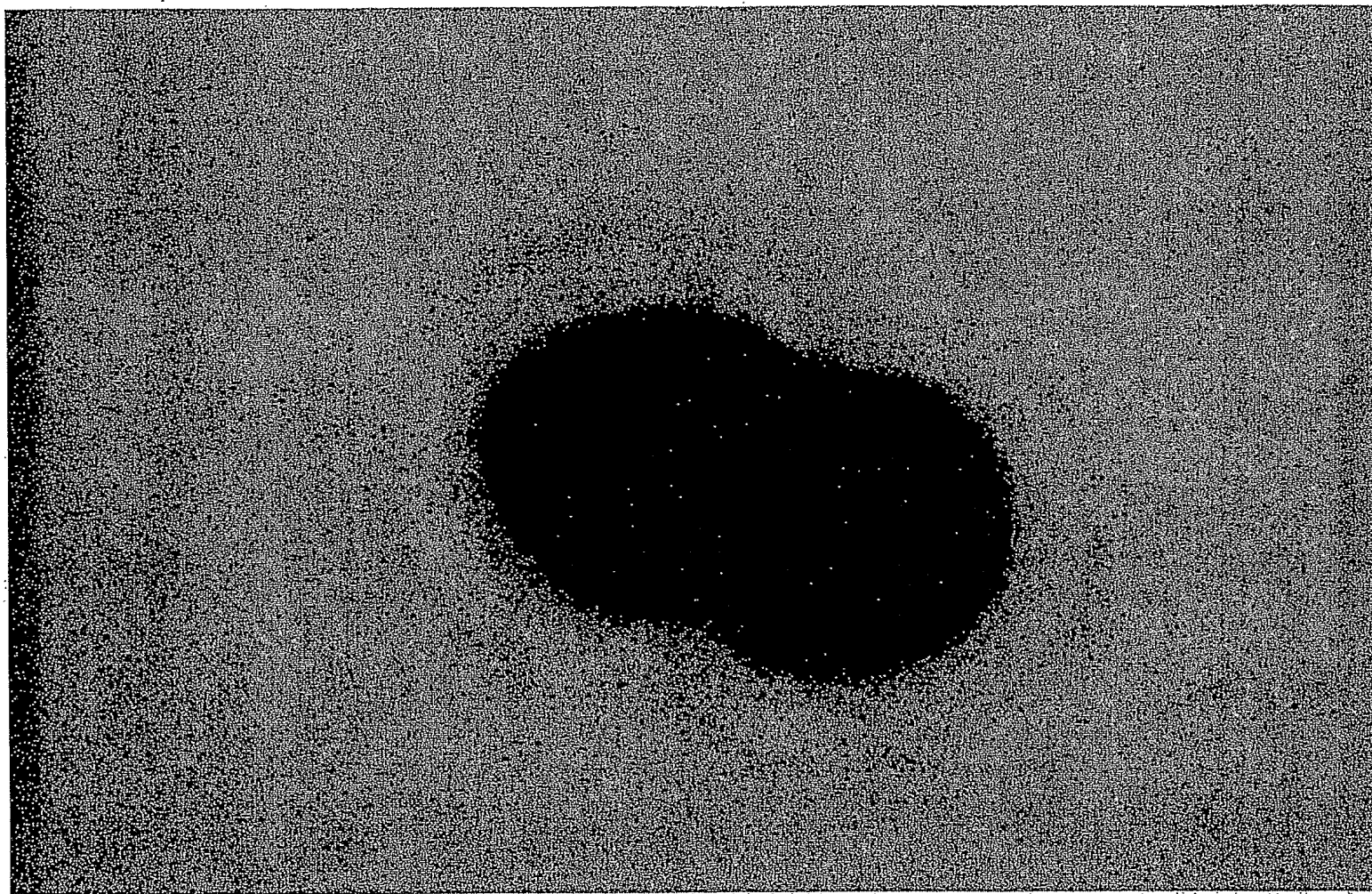


FIGURE 119

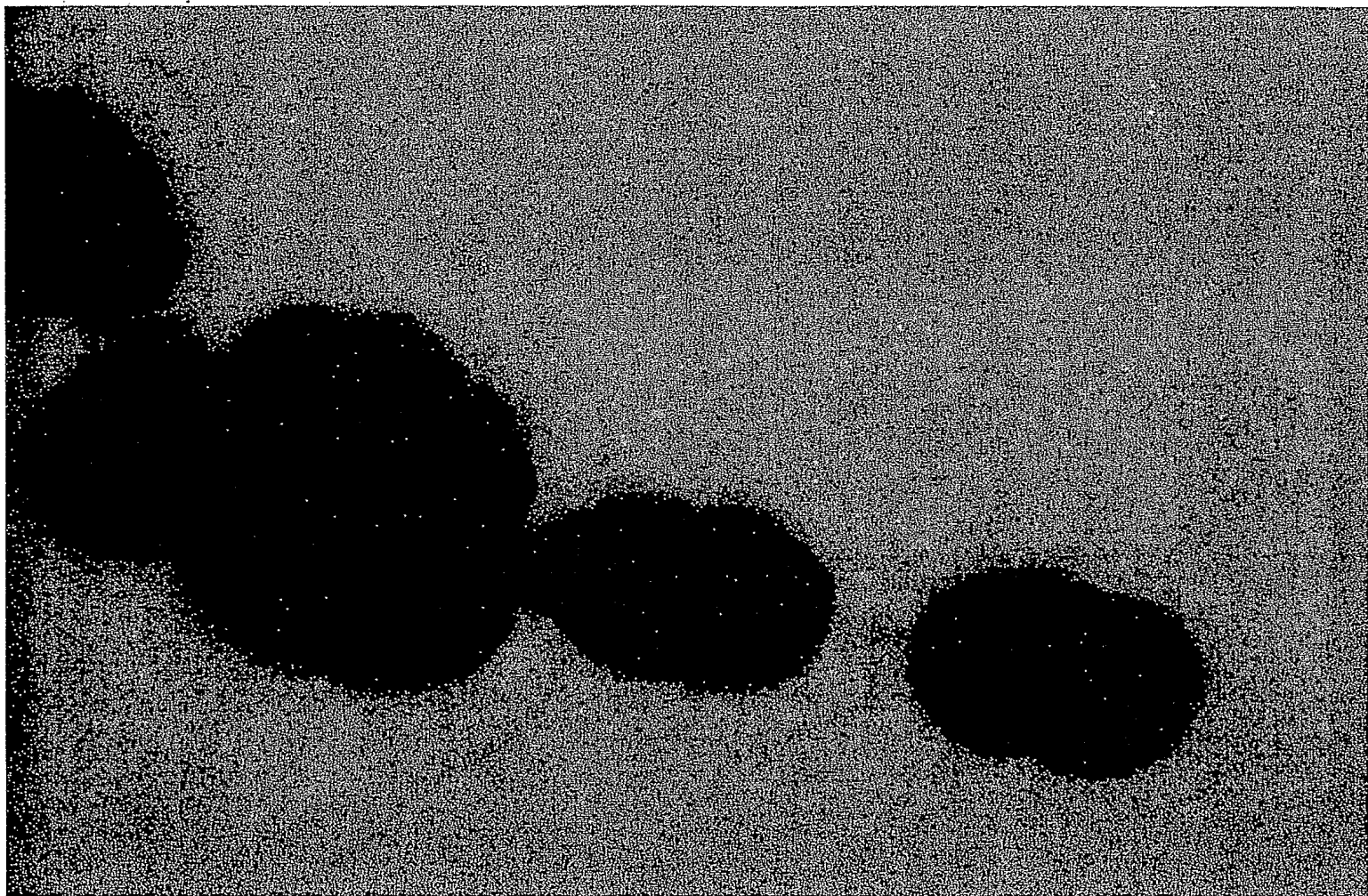


FIGURE 120

312/487

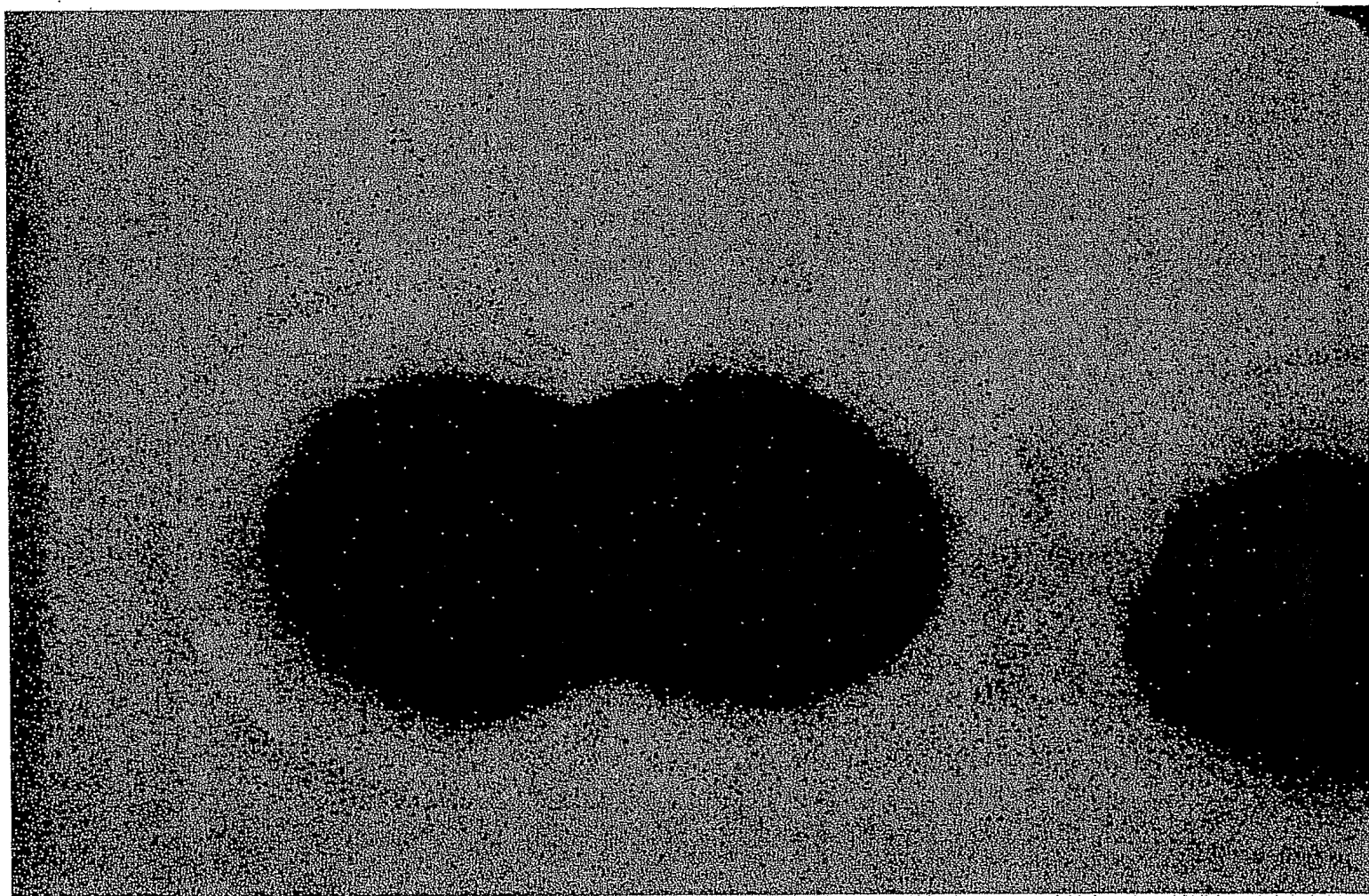


FIGURE 121

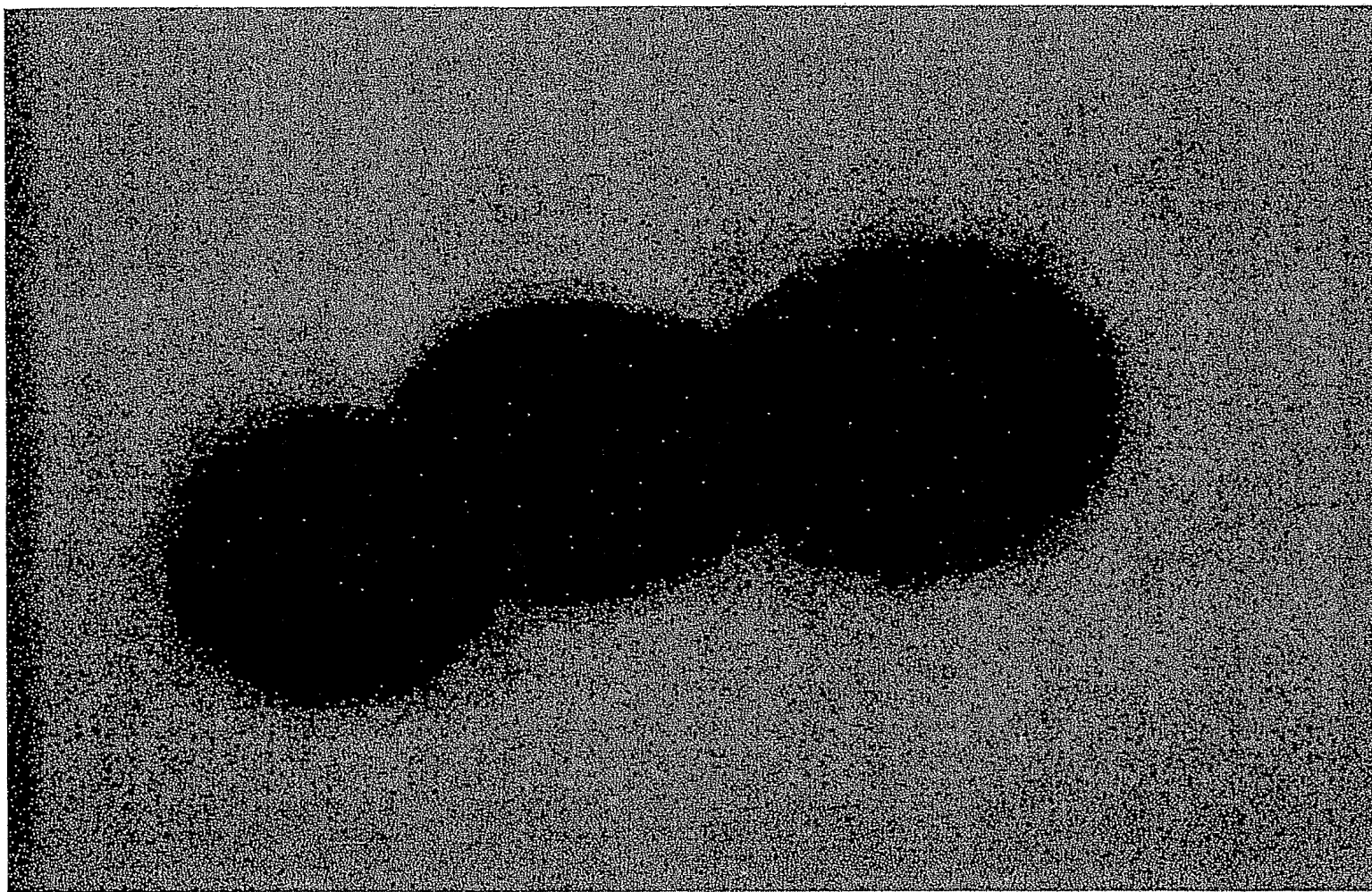


FIGURE 122

314/487

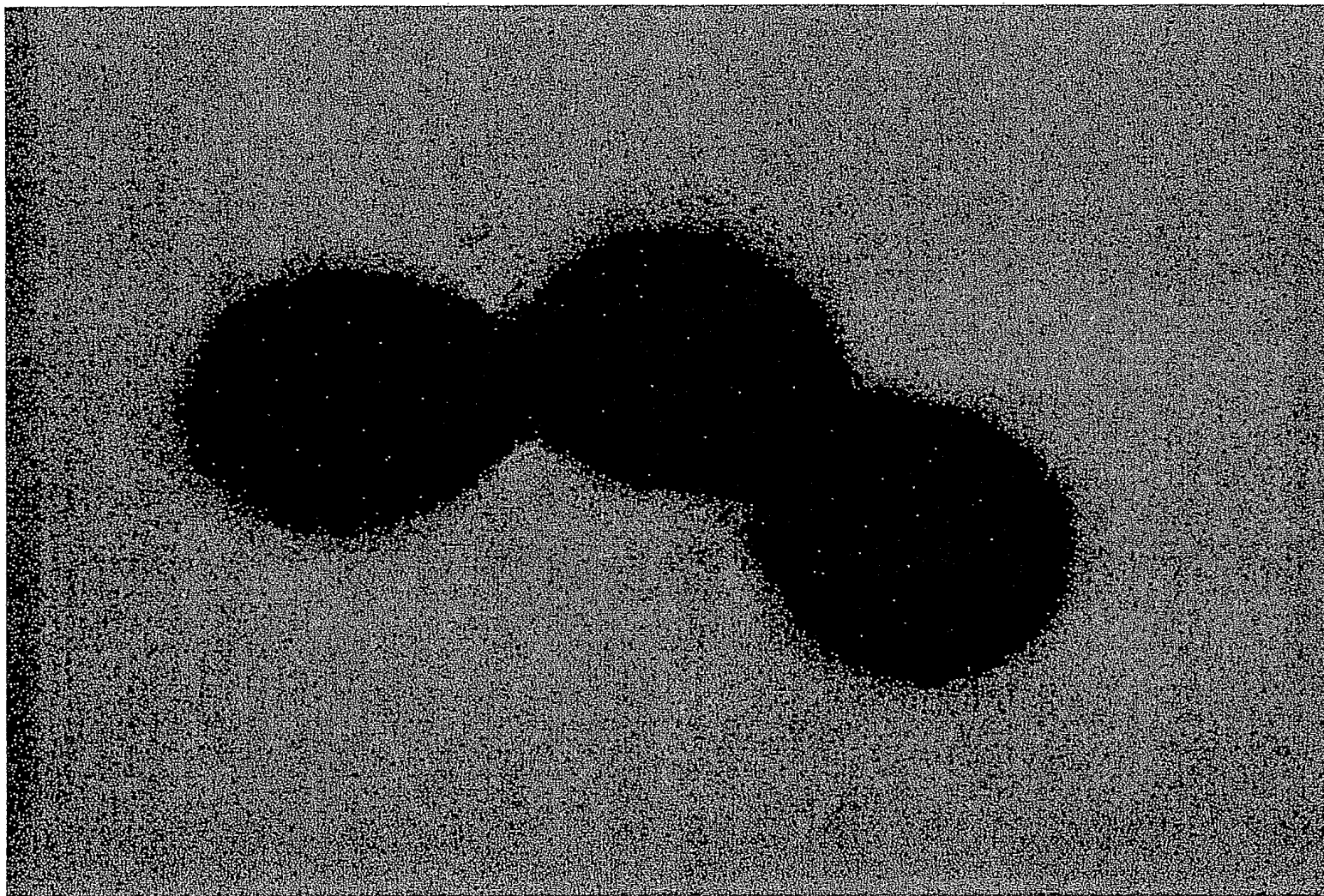


FIGURE 123

315/487

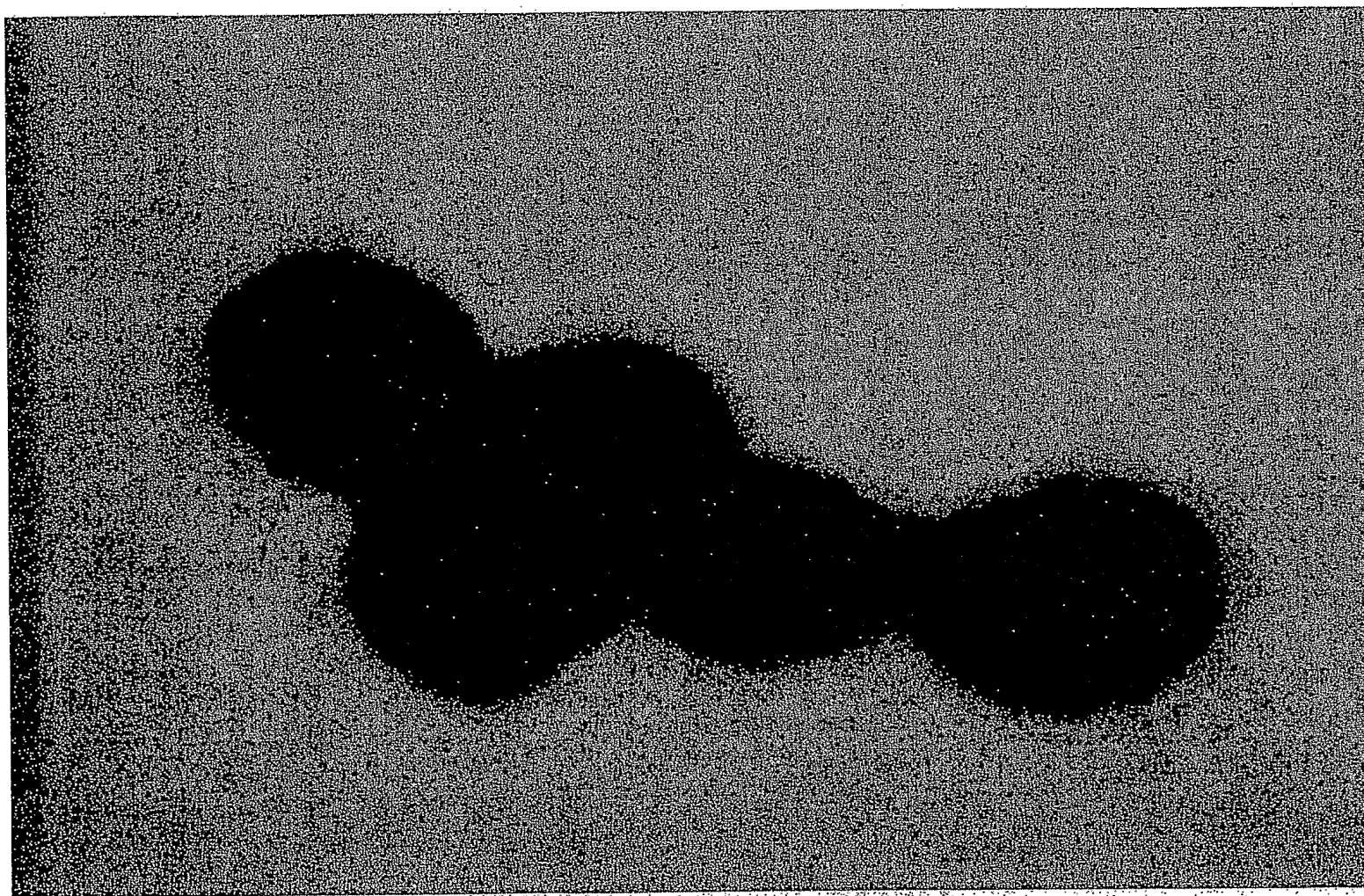


FIGURE 124

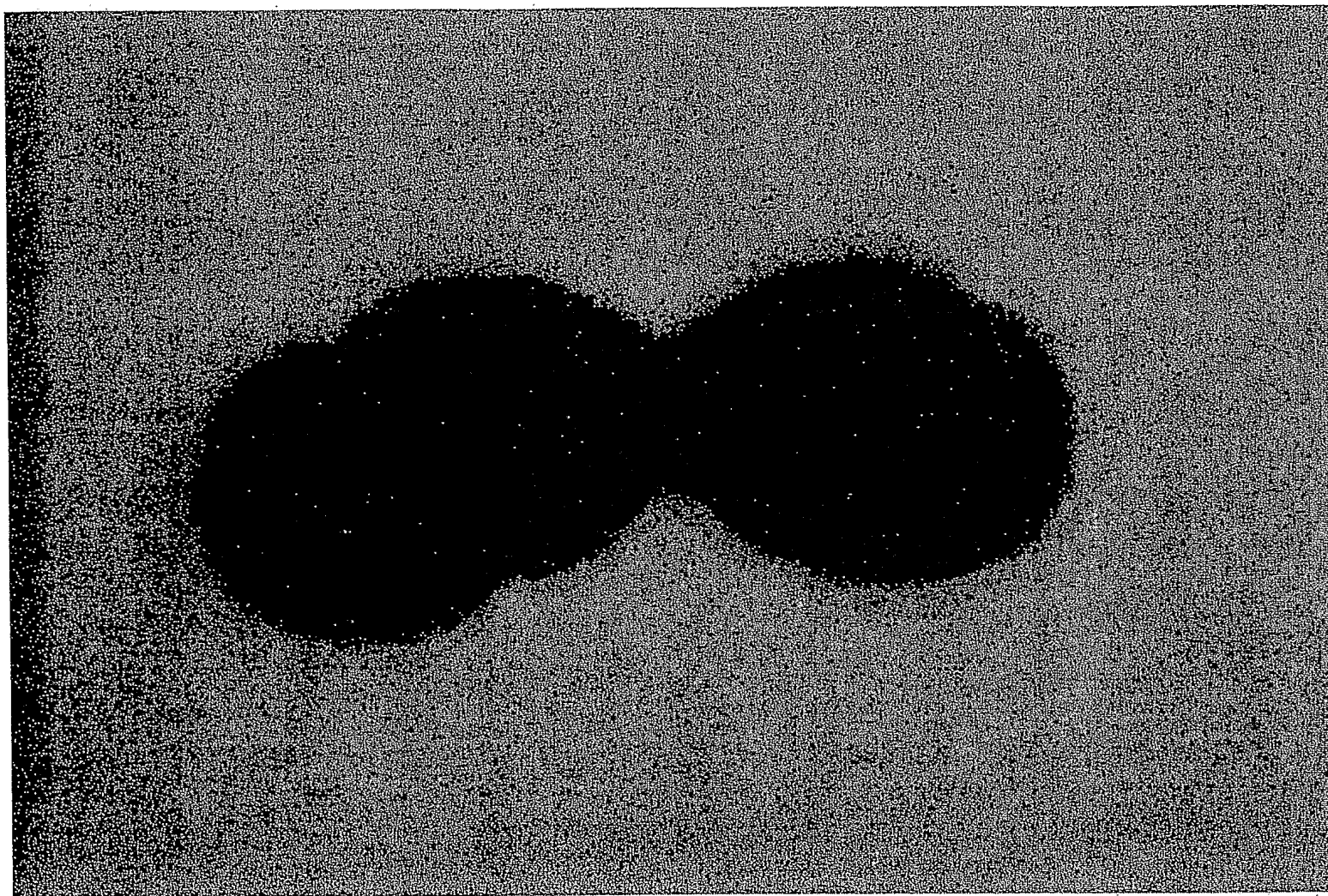


FIGURE 125

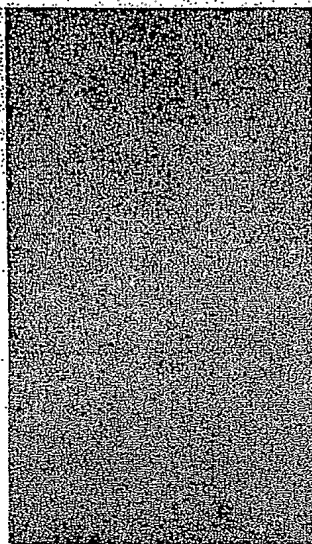
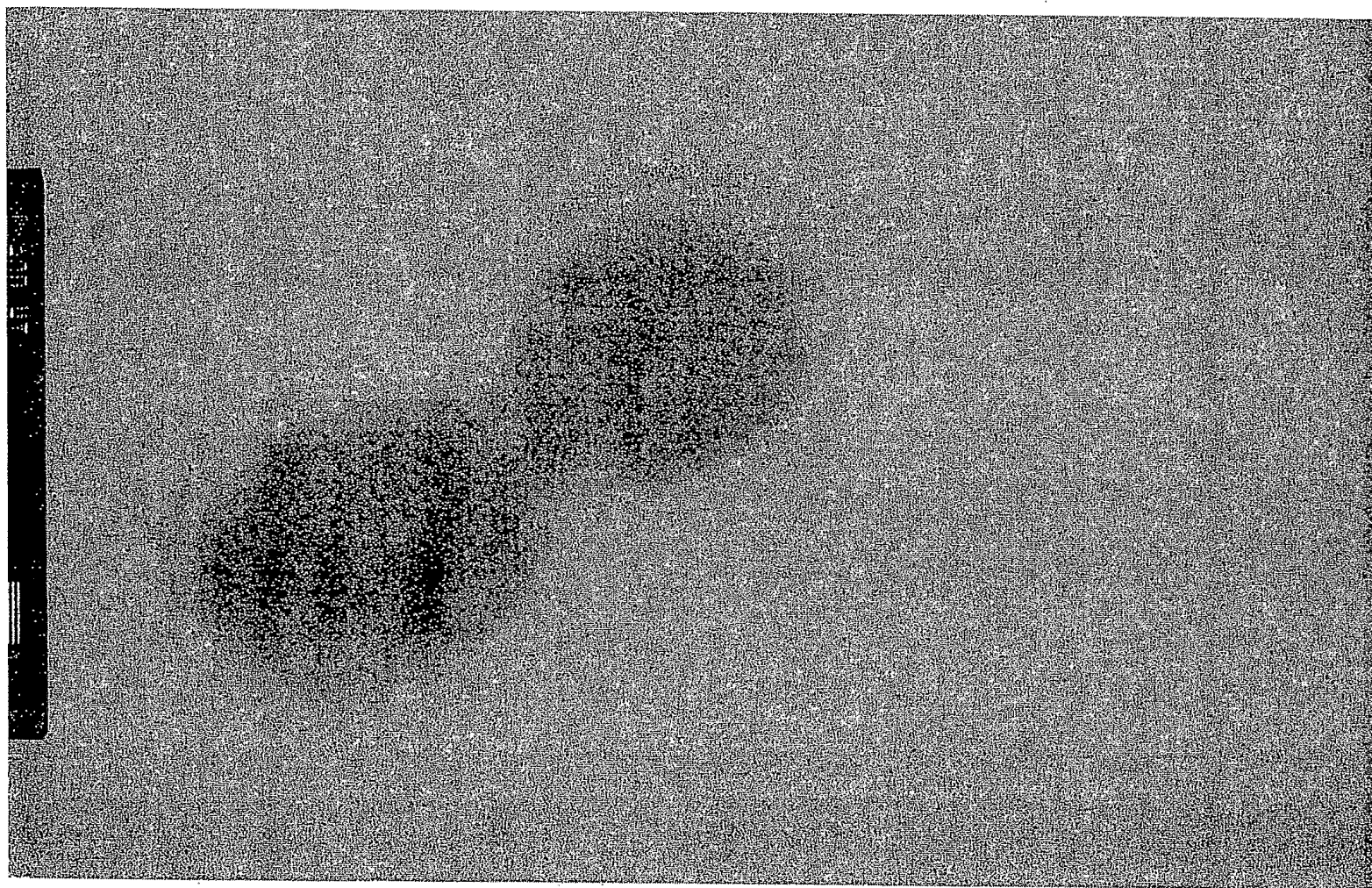


FIGURE 126

Figure 127



PCT/US05/27239 319/487

Figure 128

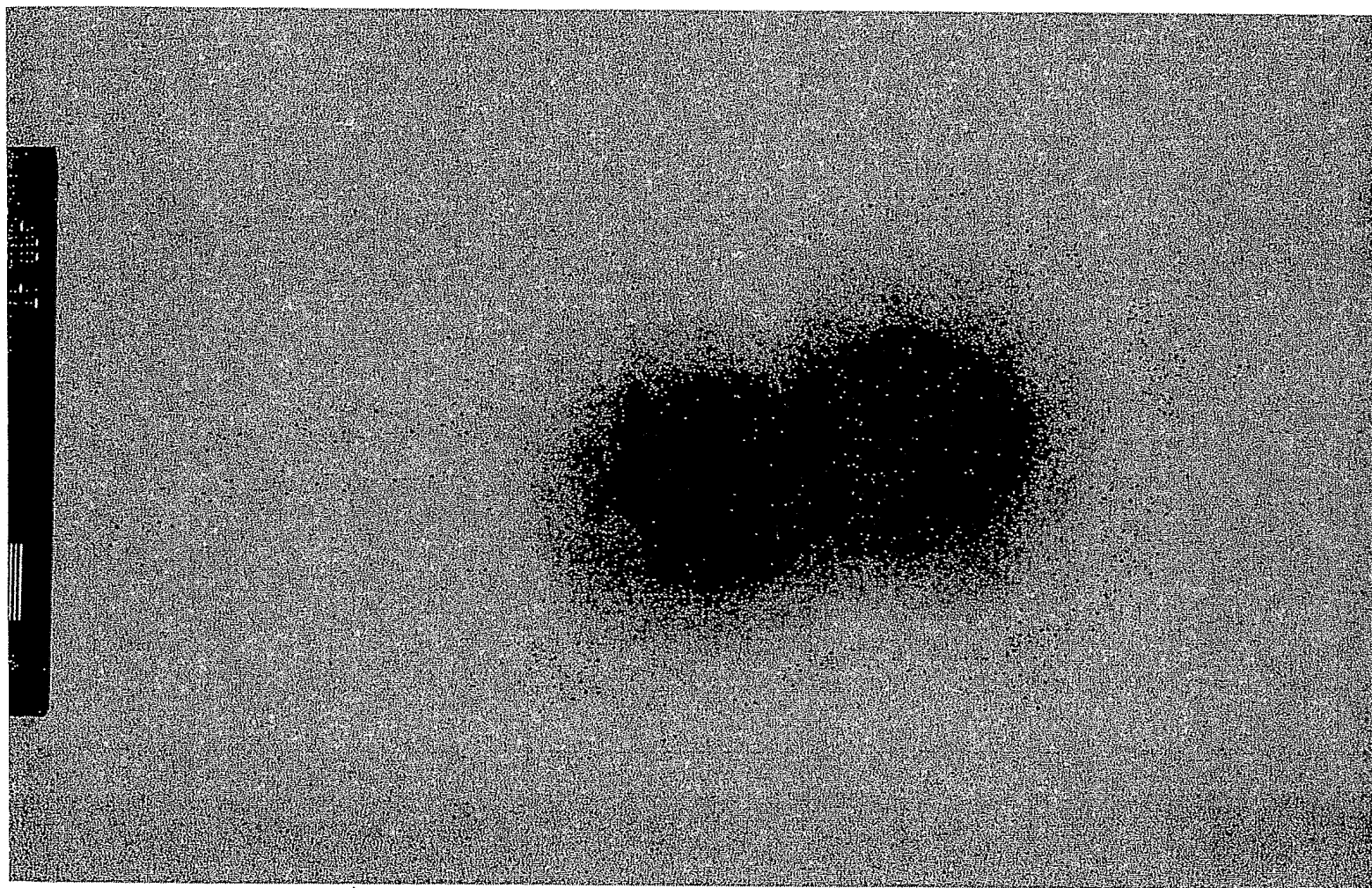


Figure 129

